



University of
Portsmouth

Department of Psychology

Factors Affecting the Accurate Detection of Trait Aggression from Gait

Liam Paul Satchell

The thesis is submitted in partial fulfilment of the requirements for the award of the degree
of Doctor of Philosophy of the University of Portsmouth.

31st March 2016

Supervised by:

Dr Paul Morris

Dr Lucy Akehurst

Dr Claire Nee

“some of the most pervasive features of my attitude to the world and to others is encoded in the way I project myself in public space”

Charles Taylor

General Abstract

This thesis extends the limited extant research on interpersonal threat detection accuracy. Broadly speaking, it was found that when viewing video presentations of walking targets, observers could accurately detect the trait aggression of those targets. The thesis opens with a summary of the relevant background literature and the general foci of this programme of research (Chapter 1). Chapter 2 explains *why* movement is important for recognising aggression. Using biomechanical analyses of gait, Chapter 2 provides evidence that individuals' dispositional aggression is related to how they walk, thus highlighting the importance of gait in detecting potential aggressors. The third Chapter outlines two experiments that demonstrate that: i) judgments of the threat posed by targets correlate with the self-reported trait aggression of the targets; and ii) these judgments of threat are most accurate when movement information is available. Chapter 4 attempts to answer the ever-present question throughout this thesis; what makes a good judge? In an eye tracking experiment, it was shown that more accurate judges of aggression are those who, when observing a video of a target walking, spend more time observing the body and legs (and not the face) of the targets. Chapters 5 and 6 investigate how age affects the accuracy of detecting trait aggression. Chapter 5 highlights that accuracy for recognising trait aggression is acquired with age. Participants from the age of 13 years were tested for aggression detection accuracy and the results show that accuracy greatly improves after the age of 18 years. The results of the experiment outlined in Chapter 6 illustrate that accuracy is maintained into older age, with participants up to the age of 91 years performing similarly to young adults. The seventh Chapter explores the methodological pitfalls of previous trait recognition research. The analyses presented in Chapter 7 show how the typical reporting of trait recognition studies: i) overestimates the size of accuracy effects; ii) ignores the variance in individual judge's ability to recognise traits; and iii) does not

acknowledge the influence of individual targets on overall accuracy values. The thesis closes with a general discussion (Chapter 8), overviewing the findings and implications (both applied and theoretical) of the current work, as well as directions for future work.

Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

Thesis Word count: 44,353 words

Liam Satchell

31st March 2016

Acknowledgements

I am grateful to the large amounts of support I have received from friends and family during my work on this thesis and I hope to return to pre-thesis levels of social engagement with these supportive people now that I have submitted this work.

Special thanks go to my supervisory team; Dr Paul Morris, Dr Lucy Akehurst and Dr Claire Nee. Their help has been invaluable and their unique, specialised, contributions have helped me to develop a thesis that ventures into different areas, pushing unique barriers that I would not have had the chance to do otherwise. Dr Jim Sauer, who set me on this path, also deserves special thanks. Your assistance with getting me motivated and set on the PhD path always will be greatly appreciated.

I would also like to thank those who have really inspired a different look at the subjects that I discuss throughout the thesis. The conversations I have had with Professor Alan Costall, Professor Vasu Reddy, Dr Beatriz Lopez and (of course) Dr Paul Morris have introduced me to new ways of considering psychology as a discipline. Their introduction to thinkers I had not previously been exposed to, such as Gibson, Neisser, James, Vygotsky and Funder, have expanded my horizons in theoretical thinking. So much so, that the theory and reasoning behind my thesis and my subsequent approaches to thinking and working in psychology have evolved in ways I did not expect at the start of this process.

I have been lucky to engage with many other academics over the course of my research programme through visits and conferences, which would not have been possible without the support I have had from Dr Lucy Akehurst, Professor Vasu Reddy, (and my conference buddies;) Dr Zarah Vernham, Dr Jackie Hillman, Dr Sam Mann and Dr Galit Nahari.

I also owe large parts of my data collection and public interaction experience (which has offered much insight into my research) to the support I have had from the International Centre for Research in Forensic Psychology (and especially Dr Claire Nee), the University of Portsmouth Aging Network (with special thanks to Dr Amy Drahota and Mr Derek Ward), the University of Portsmouth Education and Learning Team (with special thanks to Mrs Amanda Voar, Mrs Claudia Bradshaw and Ms Hannah Lawrence) and the University of Portsmouth Department of Sports and Exercise Science (with special thanks to Dr Chris Mills and Mr Liam O'Reilly). I also owe thanks to the research assistants who helped me throughout the process (notably Mr Harry Mayes and Ms Danielle Evans).

The work I present here is a reflection of the experiences and insight provided to me from those mentioned above (and so many more).

I thank you all so much for your help.

Dissemination

Journal Articles in the review process.

Satchell, L., Morris, P., Mills, C., O'Reilly, L., Marshman, P., & Akehurst, L. (2016). Evidence of Big Five and Aggressive Personalities in Gait Biomechanics. *Journal of Nonverbal Behavior*. doi:10.1007/s10919-016-0240-1

Satchell, L., Akehurst, L., & Morris, P. (in press) Learning to be streetwise: The acquisition of accurate judgments of aggression. *Psychiatry, Psychology and Law*.

Satchell, L., & Morris, P. (under review) *Describing individual differences in targets' salience and judges' detection of trait aggression*

Satchell, L., Morris, P., Morrison, E., & Akehurst, L. (in preparation) *The accuracy of interpersonal threat judgments for detecting trait aggression.*

Satchell, L., Nee, C., Morris, P., & Akehurst, L. (in preparation) *Don't look at the face; why the body matters for detecting trait aggression*

Satchell, L., Akehurst, L., Morris, P., & Nee, C. (in preparation) *Staying streetwise: the maintenance of accurate judgments of aggression into older age*

Conference presentations.

Satchell, L. (2015, April) *Judgments of threat as reflections of an approaching stranger's trait aggression*. Paper presented at the meeting of British Society for the Psychology of Individual Differences, York, UK.

Satchell, L. (2015, August) *Learning to be streetwise: The acquisition of accurate judgments of aggression*. Paper presented at the meeting of European Association for Psychology and Law, Nuremberg, Germany.

Satchell, L. (2015, September) *A walk around the world*. Poster presented at the Loch Lommond Symposium on Action Anticipation, Glasgow, UK.

Meeting and Seminar presentations.

Satchell, L. (2015, October) *Methods in my madness*. Presented at the first meeting of the Pretty Heavy Discussions Group, Portsmouth, UK.

Satchell, L. (2015, December) *Individual differences are really cool*. Presented at the meeting of the Talking About Research in Memory and Cognition (TARMAC) group, Portsmouth, UK.

Contents

Declaration.....	iii
Acknowledgements.....	iv
Dissemination	v
List of Tables	xi
List of Figures	xi
1. General Introduction	1
1.1 Thesis aims.....	2
1.2 Detecting aggression.....	3
1.3 Accurately detecting the traits of others	6
1.4 The potential consequences of inaccurate or ' <i>false-positive</i> ' threat judgments.....	12
1.5 Thesis outline	14
2. Evidence of Big Five and Aggressive Personalities in Gait Biomechanics.....	25
2.1 Introduction.....	27
2.2 Method.....	29
2.2.1 Participants.....	29
2.2.2 Materials.	29
2.2.3 Procedure.	29
2.2.4 Analyses.....	30
2.2.4.1 Statistical analyses.	30
2.2.4.2 Gait analyses.	31
2.3. Results.....	33
2.3.1 Sample Sex Differences in Personality and Gait.	33
2.3.2 Aggression and Motion Correlations.	33
2.3.3 Big Five Factors and Motion Correlations.....	34
2.4 Discussion.....	36
2.4.1 Conclusion.	38
3. The Accuracy of Interpersonal Threat Judgments for Detecting Trait Aggression.....	43
3.1 Introduction.....	45
3.2 Study One: The 'Masculine is Dangerous' Heuristic and the Accuracy of Threat Judgments	49
3.2.1 Method	49
3.2.1.1 Participants.....	49
3.2.1.2 Design.	49
3.2.1.3 Materials.	50
3.2.1.4 Procedure.	51
3.2.2 Results.....	51

3.2.2.1 Target Body Shape and Aggression.....	51
3.2.2.2 Influence of ‘masculine is dangerous’ heuristic.....	51
3.2.2.3 Accuracy at detecting trait aggression.	52
3.2.3 Study One Summary Discussion.....	52
3.3 Study Two: Controlling the Availability of Target Information.....	53
3.3.1 Method.....	53
3.3.1.1 Participants.....	53
3.3.1.2 Design.	53
3.3.1.3 Materials.	54
3.3.1.4 Procedure.	54
3.3.2 Results.....	55
3.3.2.1 How threat ratings varied across conditions.	55
3.3.2.2 The relationship between threat ratings received and trait aggression.....	56
3.3.3 Study Two Summary Discussion.....	57
3.4 General Discussion	57
3.4.1 Conclusion.	59
4. Don’t Look at the Face; Why the Body Matters for Detecting Trait Aggression	64
4.1 Introduction.....	66
4.2 Method.....	68
4.2.1 Participants.....	68
4.2.2 Materials.	68
4.2.2.1 Target videos.....	68
4.2.2.2 Eye tracking.	69
4.2.3 Procedure.	69
4.2.4 Data Analyses.	70
4.2.4.1 Critical areas of fixation.....	70
4.2.4.2 Accuracy analyses.....	71
4.3 Results.....	71
4.3.1 Observation style over time.	71
4.3.2 Observation style and accuracy.....	72
4.4 Discussion.....	75
4.4.1 Conclusion.	77
5. Learning to be Streetwise: The Acquisition of Accurate Judgments of Aggression	84
5.1 Introduction.....	86
5.2 Method.....	88
5.2.1 Participants.....	88
5.2.2 Materials.	88

5.2.3 Procedure.	90
5.2.4 Analyses.	90
5.3 Results.....	91
5.3.1 Ratings received by walkers.	91
5.3.2 Accuracy of participants.	93
5.4 Discussion.....	95
5.4.1 Conclusion.	97
6. Staying Streetwise: The Maintenance of Accurate Judgments of Aggression into Older Age	102
6.1 Introduction.....	104
6.2 Method.....	106
6.2.1 Participants.....	106
6.2.2 Materials.	106
6.2.3 Procedure.	107
6.2.4 Analyses.	107
6.3 Results.....	108
6.3.1 Ratings received by walkers.	108
6.3.2 Accuracy of participants.	110
6.4 Discussion.....	111
6.4.1 Conclusion.	114
7. Describing Individual Differences in Targets' Salience and Judges' Detection of Trait Aggression.....	121
7.1 Introduction.....	123
7.2 Method.....	126
7.2.1 Participants.....	126
7.2.2 Stimulus Materials.	126
7.2.3 Procedure.	126
7.3 Analysis and Results.....	127
7.3.1 The nomothetic approach; using the average rating received by the targets.....	127
7.3.2 Idiographic descriptions of judge accuracy; individual correlations.....	129
7.3.3 Idiographic descriptions of target personality salience; products standardised threat ratings and aggression.....	132
7.4. Discussion.....	135
7.4.1 Conclusion.	137
8. General Discussion.....	142
8.1 Overview of Main Aims and Findings.....	143
8.2 Theoretical Implications	144
8.2.1 Implications for personality psychology.	145
8.2.2 Implications for aggression detection research.....	147

8.3 Practical Applications	148
8.3.1 Benefits for interpersonal judgments of aggression.....	148
8.3.2 Benefits for detecting aggressors from afar.	150
8.4 Methodological Limitations.....	151
8.4.1 New stimuli, same old paradigm.....	151
8.4.2 Quantifying aggression.	154
8.5 Future Research	157
8.5.1 Exploring the context of judgments.	158
8.5.2 Exploring target characteristics.....	159
8.5.3 Exploring collaborative judgments of threats.	160
8.6 Thesis Conclusion.....	162
Appendix A. Ethical Approval for the Research in Chapters 2 & 3	170
Appendix B. Ethical Approval for the Research in Chapter 4.....	171
Appendix C. Ethical Approval for the Research in Chapter 5.....	172
Appendix D. Ethical Approval for the Research Presented in Chapter 6.	174
Appendix E. Ethical Approval for the Research Presented in Chapter 7.....	175
Appendix F. The Buss Perry Aggression Questionnaire Short Form (Bryant and Smith, 2001).	176
Appendix G. A discussion on reporting multiple correlation values.	178

List of Tables

Table 2.1. <i>The Mann Whitney U Differences Between Sexes for the Biomechanical and Psychometric Measures</i>	34
Table 2.2. <i>Spearman's Ranked Correlations Between Biomechanics and Psychometrics Across Participants, Split by Gender and Together as an Overall Sample</i>	35
Table 4.1. <i>The Average Percentage of Time Judges Spent Observing Each of the Body Areas of the Average Target (with Standard Deviation in Brackets).</i>	74
Table 4.2. <i>The Correlations Between Judge Overall Accuracy and Percentage of Time Spent Fixating on the Differing Areas of the Average Target Over the Length of the Video.</i> .	74

List of Figures

<i>Figure 2.1.</i> Placement of retroreflective markers.	30
<i>Figure 2.2.</i> A graphical representation of the Range of Motion measurements used in this experiment.	32
<i>Figure 2.3.</i> An alternative representation to that in Figure 2.2, presenting the Range of Motion measurements use in this study.	32
<i>Figure 3.1.</i> The appearance of targets in the two experiments.	56
<i>Figure 5.1</i> The relationship between average judgment of intimidation received by the walkers and the walker's trait aggression. Results are presented by age group and sex of walker.	92
<i>Figure 5.2.</i> The distribution of participant accuracy in detecting trait aggression with intimidation ratings.	94
<i>Figure 6.1</i> The relationship between average judgment of intimidation received by the walkers and the walker's trait aggression. Results are presented by age group and sex of walker.	109
<i>Figure 6.2.</i> The distribution of participant accuracy (individual correlations) for each age group.	111
<i>Figure 7.1.</i> The correlation between the average threat rating received by the target and A) the average masculinity ratings received by the target; B) the trait aggression of the target.	
<i>Figure 7.2.</i> The distribution of A) participants' association between threat and masculinity ratings and B) participant accuracy in detecting trait aggression with threat ratings.	131
<i>Figure 7.3.</i> The distribution of A) average target salience and B) variance in target salience.	134

1. General Introduction

1.1 Thesis aims

When walking down a street, an approaching stranger could be seen as a potential threat to personal safety. Many people have experienced feeling ‘threatened’ or ‘intimidated’ by other people when walking at night. Individuals may change behaviours by crossing streets or changing their own behaviours in response to someone who may pose them harm. However, the psychological processes behind feelings of threat or intimidation are poorly understood. Further, there is little psychological research that explores the ‘benefits’ of such feelings. Are such feelings, in fact, accurate reflections of the truth of the risk posed to a perceiver? If so, how can judgments of threat reflect the disposition of another person from simply seeing that person walking towards you?

The thesis investigates the accuracy of threat and intimidation judgments at predicting the dispositional aggression of an approaching person. To do this, seven studies (organised into six chapters) were designed to explore and improve upon the methodological, theoretical and analytical understanding of threat judgments. These studies aim to improve on the methodology of the extant work in threat detection research, using realistic presentations of other people. Most of the research conducted in interpersonal threat judgments is reliant on making judgments of photographs of faces (see Section 1.2). In the studies presented here, the selection of stimulus presentation is based on the context being investigated – judging the threat posed by an approaching person. Thus realistic presentations of approaching people are used throughout the thesis. Another aim of the thesis is to adopt and thoroughly investigate the Realistic Accuracy Model (RAM, Funder, 1999) of personality judgment accuracy. This model proposes that judgments of another’s traits are accurate when a target’s behaviours, *relevant* to the trait being judged (see Chapter 2), are *available* (see Chapter 3) for a judge to *detect* (see Chapter 4) and correctly *utilize* (see Chapters 5 & 6). It should be noted that, to date,

aspects of the RAM (beyond *utilization*) have not received detailed investigation (see Section 1.3) and it is one aim of this thesis to expand the theoretical understanding of the RAM. The third main aim of the thesis is to emphasise the importance of an individual differences approach to analysing judge and target influences on accuracy in threat judgments. There is very little research which reports the accuracy of individual judges in detecting traits and no known literature that details variation in the salience of a target's traits. Here, the ability of individual judges to detect trait aggression is considered throughout and the thesis concludes with a commentary on variation in the salience of a target's aggression (Chapter 7).

We know little about the accuracy of threat judgments in detecting the traits of approaching people. It is the purpose of this thesis to provide important methodological, theoretical and analytical advice to future research in trait judgments and for settings where threat judgments would be of the utmost importance.

1.2 Detecting aggression

Judgments of threat should be considered accurate when they are able to detect the dispositional aggression of another person. The majority of existing research investigating judgments of threat does not test for accuracy, instead reporting perceptions of what is threatening alone. To best understand the value of threat judgments, it is important to understand if judgments of threat can accurately reflect any genuine malevolent tendencies of a target. This is, perhaps, because the existing literature is dominated by evolutionary models of threat judgment which prioritise making false positive, 'safe choices' of danger. Accuracy is not an important feature of a theoretical framework where humans are expected to see threats where there are none. What follows is a literature review of the existing research into judgments of threat and an overview of the frequently cited Error Management Theory.

In wider threat judgment research, psychologists have investigated perception of potential danger from photos of animals (such as snakes and spiders, Lipp, Derakshan, Waters & Logies, 2004; LoBue & DeLoache, 2009; Öhman, Flykt & Esteves, 2001), dangerous objects (such as guns and syringes; Blanchette, 2006; Fox Griggs & Mouchlianitis, 2007) and human faces (Boshyan, Zebrowitz, Franklin, McCormick and Carré, 2013; Carré, McCormick & Mondloch, 2009; Geniole, Molnar, Carré & McCormick, 2014; Hehman, Flake & Freeman, 2015; Hehman, Leitner & Gaertner, 2013). It is a robust finding across the literature that stimuli, which pose more of a threat, are more readily recognised than those which do not. This prioritised identification of things that are potentially harmful could be an important adaption for survival; quickly detecting dangers allows more time for avoidance behaviours. However, all of the aforementioned research focuses on brief presentations of stimulus photographs. Some authors of research into danger detection from faces (e.g. Geniole, et al., 2014; Hehman, et al., 2015; Hehman, et al., 2013) consider the presentation of photographs a strength as, they argue, it shows that relatively little information (static, select presentations) can communicate potential danger. Whilst it is impressive that judgments from static imagery are accurate, it may well be the case that interpersonal judgments adapt and change over time, beyond, for example, the first 39 microseconds of observation of a target (such as used by; Carré, et al., 2009). These judgments may further vary in accuracy when the full image of a target person is available and that person is in motion (as would usually be the case in everyday life). It is exactly this concept that the thesis investigates. All the judgment studies in this research programme use dynamic stimuli where gait is clearly available to the participant. All the targets are filmed walking towards the camera, as if they were to be approaching the judges. Whilst this experimental setting is still not typical of making judgments of threat in

everyday life, the choice of stimulus presentation is a move towards a meaningful everyday context for judgments of danger.

Research on intra-human threat detection often uses evolutionary theory to explain why and how judgments of threat occur. For example, Schaller, Faulkner, Park, Neuberg and Kenrick (2004) propose that perceptions of the threat a stranger poses are the product of the tribal evolutionary history of human beings. Historically human beings lived in small groups or tribes, and members of other tribes posed a real threat to the wellbeing of one's own tribe. It is the argument of Schaller et al. (2004) that judgments of other people as threatening is a logical extension of this survival technique. There is also neurological evidence of bias towards seeing danger where there is none, notably in Damasio's (1994) somatic marker hypothesis. This theory, like Error Management Theory, is based on the idea that judgments of the threat posed by the environment are emotionally biased towards more danger than there is (Dunn, Dalgeish & Lawrence, 2006). However, adopting this view point of 'anything atypical to one's environment could be a danger' could lead to the needless use of avoidance or cautious approach tactics. Hasleton and Buss' (2000; 2009) popular Error Management Theory proposes that individuals are evolutionarily prone to over-estimation of risk; assuming there are dangers where there are none. Error Management Theory has been widely used to explain judgments of danger, including the decision of fire arms police officers to shoot a potential threat (Plant, Goplen & Kunstman, 2011). When judgments of danger are being made in urgency, there are benefits to erring on the side of caution and using simple heuristics to judge threats (Gigerenzer & Gaissmaier, 2011). However, in situations where one maintains an effective defensive distance (see; McNaughton & Corr, 2004) it is likely that accurate judgments of danger can be made.

Recent research has shown that participants can accurately judge a stranger's fighting ability (Sell, et al., 2009), from observing at a photograph of a face. Furthermore research into the accurate identification of less malevolent traits than threat or aggression (for example, social skills and creativity) frequently demonstrates that accurate identification is possible from surprisingly little information (such as, Albright, Kenny & Malloy, 1988; Funder, 1980; Vazire; 2010; Section 1.2 below). Given that this research shows that accuracy in detecting someone else's disposition is possible, it is somewhat surprising that research into judgments of interpersonal threat are not routinely investigated in terms of accuracy.

Judgments of threat can be costly if they are not accurate. It should be of the utmost importance to understand who appears threatening and if they are truly a threat. It should be of great interest to understand who is a 'better' judge of threat and what features of that judge make them better. This thesis is an attempt to better address these concepts and to try and explain the accuracy of interpersonal threat detection.

1.3 Accurately detecting the traits of others

It is possible that judgments of threat can relate to the trait aggression of an approaching person. Considering accuracy in threat judgments as simple recognition of a target's traits is similar to the existing body of literature on judgments of personality traits. Research has demonstrated that accurate judgments of personality traits of unknown people is possible from brief interactions (Albright, et al., 1988; Back, Schmukle & Egloff, 2010; Funder, 1980; Kenny & Albright, 1987; Kenny, Albright, Malloy & Kashy, 1994; Levesque & Kenny, 1993; McCrae, 1982; Vazire; 2010). There is even evidence that accurate judgements about traits can be made from short excerpts of speech (Hu, Wang, Short & Fu, 2012) or photographs of faces (Carré et al, 2009). This research shows that

judgments of another person's personality can be accurate, but accurately judging the qualities of an approaching person from a distance is less well understood.

In a real world context, judgments of another person are not first made when a judge is interacting with another person, hearing the way that person talks or observing the shape of that person's face. However, as that target approaches the perceiver, that target's gait would be available to the judge, when perhaps other information is not. It is possible that accurate judgments of that target person could be formed from this gait information. In fact, this is beneficial to judgments of trait aggression, as judgments of threat made from a distance could allow that judge to act before they are in close proximity to the target. There is evidence of animals attempting to maintain an effective 'defensive distance' (McNaughton & Corr, 2004), where the intensity of a potential danger affects how the animal acts in proximity to the danger. It is proposed that the animal makes the judgment to approach (continue towards and prepare to defend itself) or avoid (to run away from) the potential danger. In interpersonal threat judgments, accurate recognition of a target's trait aggression at a distance would allow humans to engage in similar approach or avoidance behaviour. It is possible that gait is the best source of trait information when at a distance.

There are many studies that investigate judgements about other people based solely on their gait. For example, studies have demonstrated that gait can be used to make a range of accurate judgments, including the recognition of the sex of a walker, (Johnson & Tassinari, 2005; Kozlowski & Cutting, 1977), or even identify a known person (Cutting & Kozlowski, 1977). It has even been shown that gait can be used to inform judgments of vulnerability from simple point light presentations (Gunns, Johnston & Hudson, 2002; Johnston, 2012). Experiments investigating perceptions of gait or movement alone often use 'point light' presentations of their stimuli. Point light displays were developed by Johansson (1973) to present a body as simple dots outlining the key parts of a person.

There are few studies that have investigated judgements about personality traits based on gait. Thorsen, Vuong and Atkinson (2012) found that participants made reliable but not accurate judgments of the Big Five traits of point light walkers. The reliability but not accuracy of the judgments is interesting, as it suggests that participants had a collective expectation of a trait from a gait style but this did not match the actual trait characteristics of the stimuli. It could be the case that using walkers represented in point lights misrepresented the actual gait of the walkers, hence the discrepancy between participant consensus and target traits. This consensus in participants' judgment is important though, as it suggests that participants are interpreting the gait of the stimulus person in the same way. Perhaps, if participants observe gait as part of the presentation of a fully visible person (not a point light display) gait would provide more veridical information about the walkers.

Funder (1995; 1999) proposed a model demonstrating how accurate recognition of traits in other people is possible. Funder emphasises the importance of considering features of the judge and the target in personality judgment in this model. It could be argued that most existing research on personality judgments considers the judge as the only source of variation when testing for trait recognition. It would be truer, in interpersonal perception research, to consider the information available from the targets as equally important. The judge can only judge the traits of a target when that target's traits are made obvious to the judge. To this end, Funder (1995; 1999) proposed a four-step Realistic Accuracy Model (RAM) to emphasise the importance of the information available to the judge from the target person. The RAM suggests that an accurate judgment of a target's traits can be made if the target reveals trait *relevant* information that is *available* for the judge. These first two features of the RAM highlight that the target is important to the judgment of traits as well as the judge. Accurate judgments of another's traits can only be possible if the target's

properties relate to their traits (e.g. if gait related to trait aggression) and the target was displaying that property (e.g. the target is walking). Thus, it is only if the information from the target is *available* and *relevant* then it is possible that judges could make accurate judgments of traits. To achieve accuracy judges must still successfully *detect* and then *utilize* the target's information. The third and fourth stage of the RAM explains how the judges could still be inaccurate at recognising the traits of a target even when the relevant targets' traits are available. The RAM emphasises that a judge must detect the relevant information from the target (to continue the above example, e.g. if the judge observes the gait of the target and not simply the target's face) and then utilize (e.g. know that the gait being observed is aggressive) that information to reach a judgment. Detection and utilization are difficult to separate experimentally, and Funder notes: "the distinction between detection and [...] utilization, is the same as the distinction between perception and cognition, and it is just as blurry." (Funder, 1999, p.131). By considering the qualities of both target and judge, the RAM offers a novel approach to understanding how judgments of threat may be accurate.

One of the main aims of this thesis is to use the RAM as a framework for the experimental work but also to expand the theoretical premise of Funder's work. Whilst the RAM offers a highly practical view of how trait identification accuracy may be possible, there are theories of interpersonal information that Funder, perhaps wrongly, discounts. This includes the works of J.J. Gibson and William James (Funder, 1999). The idea that the world (and implicitly another person) has 'affordances' (information based on behavioural, psychological and physical interaction between a perceiver and the environment) to an active perceiver lends itself nicely to explaining how accurate detection

of others' traits is possible¹. Despite the theoretical utility of adopting an approach to social psychology based on interaction (an 'ecological' approach), "a fully ecological or mutualist social psychology has failed to develop" (Good, 2007, p285). Funder is critical of Gibson's model of *affordances* in the context of personality judgment and provides two main arguments for his critical view.

Funder's first critique is driven by the fact that there is evidence that judges are more accurate at discerning the 'general' traits of an individual rather than 'domain specific' behaviours. For example, Funder notes "people are better at judging another person's general degree of talkativeness than at judging how talkative he or she will be specifically with them" (1999, p.85). Funder implies that personality affordances would lead to accurate, domain specific, judgments of traits. Funder's critique is somewhat counter to what one would suggest from a theory of interpersonal affordances. In fact, a theory of interpersonal affordances would argue that *a priori* predictions of how an individual will act in relation to a perceiver have no reason to be accurate. Accurately predicting qualities of a future interaction with another person would require previous exposure to that person's social affordances. More than anything else the theory of affordances requires *active perceivers*, and so the perceiver would have to be act towards and react to the target to elicit the relevant interpersonal affordances. However, an individual's general disposition will be related to how an individual interacts with others, moves through space and is generally presented, thus offering more general affordances to a perceiver (a vicarious active perceiver in this context.) Ergo, from a theory of social

¹ It should be noted that Gibson was not in favour of a single theory being used to explain social and asocial perception (Gibson, 1951; see Costall, 1995 for a review). In fact it was largely the work of William Swann (1984) and colleagues that proposed the utility of social affordances. Further, some of these proposed adaptations of the theory of affordances do not adhere to the importance of an active perceiver (see Zebrowtiz & Collins, 1997) and draw parallels to evolutionary theories of psychology (see Zebrowtiz & Montepare, 2005) which do a disservice to the original theory of affordances (Gibson rejected such nativism; Costall & Morris, 2015).

affordances we would expect predictions of broad behaviours (such as general talkativeness) to be accurate as the judge has access to broad affordances. Judgments of specific dyadic behaviours with a judge (such as ‘how talkative he or she will be specifically with them’) would require interpersonal affordances relevant to that interaction. If a judge does not have access to the relevant affordances, the judge cannot be accurate in predicting behaviours.

Secondly, Funder (1999) argues that perception is possible without the perceiver being *active* during judgment. Many research paradigms utilised for interpersonal perception experiments (including the current PhD programme) require participants to make judgments about complex qualities of a target based on asocial and fundamentally impoverished presentations of target stimuli. Often judgments produced in these paradigms are accurate predictors of a target’s traits (such as; Carré, et al., 2009). However, Funder’s scepticism of the need for active perceivers is driven by the assumptions of laboratory research². A rather obvious point is the fact that participants in psychology experiments have engaged with other people in their lifetime before entering the laboratory. They have been exposed to an array of broad interpersonal affordances from individuals with whom they have engaged, lived, worked, co-operated and fought throughout their lives. A wide variety of (dyadic) interactions have allowed perceivers the information necessary to extrapolate accurate judgments from asocial materials during experiments. No participant arrives at the laboratory a clean state. The participant’s wealth of social knowledge outside the world of the experiment is likely to influence their performance in the laboratory.

² It is important to note that persistence and justification of ‘passive observer’ paradigms is most likely driven by the absence of a popular, alternative methodology. Experiments with ‘active perceivers’ would require a break from the tradition of vigorous experimental control (or rather, continuity) and ready separation of stimulus and perceiver. As Good notes: “Laboratory psychology has a built-in resistance to mutualism” (2007, p.284)

In this thesis, I largely adopt Funder's (1995; 1999; 2012) approach to understanding trait judgment accuracy, as it does the rare service of highlighting the importance of the target as part of the judgment (in Funder's terms the *relevance* and *availability* of a target's actions to their traits). Largely speaking the whole RAM has not been thoroughly explored, with a clear lack of research on the *relevance*, *availability* and *detection* aspects of the model. The thesis expands the RAM using evidence from all aspects of the model. More than this, from a theoretical perspective, I go further than Funder and incorporate concepts from the theory of affordances (Gibson, 1979) including its more recent commentaries (Costall, 1995) to explain my research findings.

Funder's RAM observes that different perceivers may utilize information from a target in different ways. The features of a good 'judge' of traits are largely unexplored in the literature³ but it is perfectly plausible to expect judge ability to vary. Costall (1995) notes that the understanding of affordances (for this thesis, interpersonal affordances in particular) may be culturally acquired; "a child, for example, is not simply left to 'discover' the function of a cup or spoon; rather the learning situation involves careful structuring" (p.472). I suggest that Costall's comments on the social and cultural understanding of affordances from objects can also be used to understand age and cultural effects on the utilization of interpersonal affordances in the RAM. Just as a child learns the meaning of a spoon through its social context, so might the child learn to recognise threatening gait⁴. This, theoretically, is how individuals might acquire and perfect accurate judgments of the aggression of others through experience and structured environments.

1.4 The potential consequences of inaccurate or 'false-positive' threat judgments

³ This is in large part due to perceived methodological restrictions by many of the extant studies in this topic, see Chapter 7

⁴ See Chapter 5

Poor judgement of the potential threat posed by another person could be detrimental to the perceiver. Assuming someone is not a risk when, in fact, they are could lead to injury or worse. Assuming an approaching person is dangerous when they are not could lead to self-inhibiting and anxiety-related behaviours. A growing body of literature has demonstrated that how much an individual fears crime in their local area effects the time they spend walking (Foster, Knuiman, Hooper, Christian & Giles-Corti, 2014a), cycling (Kramer, Maas, Wingen & Kunst, 2013) and engaging in healthy activity programs (Dawson, Hillsdon, Boller & Foster, 2007). If someone cannot (or believes they cannot) accurately discriminate between more or less aggressive individuals, they may feel more afraid of being a victim of crime in general.

In the above examples adults are restricting their own behaviours because of the perceived risk of crime. A further problem develops when an individual's behaviours (and therefore potential learning experiences) are inhibited by others. Over the last two decades, children have become increasingly protected from making their own judgments of danger (Foster, Villanueva, Wood, Christian, & Giles-Corti, 2014b), which in turn could undermine their acquisition of threat detection accuracy through experience. This is in large part due to the increase in parents' fear of strangers which has led to more supervision of their children (e.g. on the walk to and from school; Foster et al., 2014b). Without the opportunity to be an active participant in judging danger it is possible that the ability to acquire accurate threat detection (and by extension have confidence in one's ability to detect threat) may be impeded. It is therefore important to study the accuracy of interpersonal risk judgments in the most 'at risk' populations. This thesis fills this gap in the literature by exploring the threat detection abilities of children (see Chapter 5) and older adults (see Chapter 6). Inferences from these studies could help to structure an

understanding of how accurate judgments of danger are acquired and whether they are maintained.

1.5 Thesis outline

In the PhD, I consider how judgments of the *threat*⁵ an individual poses may relate to the trait aggression of that individual. Over the course of six studies I address the degree to which it is possible for interpersonal judgments of an approaching person to reflect that person's trait aggression.

Chapter 2: Evidence of big five and aggressive personalities in gait. If gait information is critical to accurate human judgments of aggression, then what biomechanical aspects of gait are communicating dispositional aggression? In fact, in Chapter 2 it is more generally asked whether or not individual differences in gait relate to individual differences in personality, including the Big Five model of personality. The gait of 29 participants was motion captured and then critical aspects of gait (the thorax range of movement [ROM], the pelvis ROM and the thorax ROM relative to the pelvis ROM) were correlated with the participants' self-reported aggressive (physical and verbal aggression and trait anger and hostility) and Big Five (extraversion, conscientiousness, neuroticism, agreeableness and openness to experience) personalities. It was shown that there were multiple relationships between gait and personality, but critically Physical Aggression (the self-report measure of aggression used in all the experiments of this thesis) correlated with the relative thorax to pelvis ROM. This aspect of gait is used to explain how judges in the other experiments presented in this thesis are able to detect trait aggression from observing walking targets.

⁵ In the studies presented in Chapters 5 and 6, participants made ratings of *intimidation* rather than threat. Intimidation was considered, a more relatable and more routinely experienced construct for our younger and older participants.

Chapter 3: The accuracy of interpersonal threat judgments in detecting trait aggression. This chapter establishes the core topic of research for the thesis, demonstrating that judgments of threat do accurately reflect trait aggression of a target and that gait is an important factor in judging trait aggression. In the two experiments presented in this chapter undergraduate students made judgments of the threat posed by target stimuli. In the first study 61 participants judged the threat posed by videos of 23 targets walking on a treadmill, facing the camera. From these judgments it was evident that judgments of threat could accurately discern between more and less aggressive targets. Furthermore, this accuracy was not facilitated through the body shape of the targets. To explain why accuracy was possible, a second experiment was conducted. Participants viewed static information alone (screen shots from videos of the targets, $N = 20$) dynamic information alone (point light motion capture displays of the targets, $N=20$) or full information (the videos of the stimuli that were viewed in the first experiment, $N=19$). It was found that, accuracy was greatest when participants had access to dynamic information. This chapter concludes by highlighting the importance of gait information when judging the aggression others.

Chapter 4: Don't look at the face; why the body matters to detecting trait aggression. Given that gait style is important when communicating aggression (Chapter 2), it could be the case that more accurate participants are simply those who prioritise gait when observing a target. In this experiment, 27 participants watched full videos of 12 targets walking and made judgments of how threatening they perceived the targets to be. All participants wore eye tracking equipment. It was found that participants' gaze was directed at different areas of the targets over the 10 seconds that each target was visible. The majority of participants focussed on the face of targets in the first second of the footage. However, after the first second, an interesting finding emerged. Participants who

were more accurate at detecting trait aggression (across the 12 targets) fixated more on the torso and legs of the targets compared to fixations on their heads. Those who spent more time observing the heads of the targets were less likely to be accurate. In simple terms; observation of gait-relevant features of the target stimuli led to more accurate judgments of aggression. This, in effect, is a naturalistic replication of the findings of Chapter 3's second study; accuracy was greatest when dynamic information was prioritised by participants.

Chapter 5: Learning to be streetwise: The acquisition of accurate judgments of aggression. It is demonstrated in Chapter 3 that accurate judgments of others' aggression is possible from simply watching that person walking (i.e. via point light displays of target stimuli). It is not known, however, when this ability to detect accurate dangers from others develops. This chapter explores the accuracy of 13 to 15 year olds ($N=85$), 16-17 year olds ($N=103$) and over 18 year olds ($N=54$) when observing point light presentations of nine people walking. It is notable that the results of the over 18 years old group replicated the accuracy findings presented in Chapter 3. When comparing the age groups, it was found that the over 18 year olds were the most accurate but there was no difference in accuracy between the younger two age groups. It was further demonstrated that consensus in ratings made by participants improved with age; the oldest participants showed greater agreement across ratings than the younger participants. The results presented in this chapter demonstrate the acquisition of trait aggression detection accuracy with age and the sudden improvement in accuracy after the age of 18 years.

Chapter 6: Staying streetwise: The maintenance of accurate judgments of aggression over age. Given that Chapter 5 presents evidence showing that accuracy improves from a young age, it was of interest to know whether accuracy is maintained into older ages. In this chapter, participants watched full videos of nine people walking and made judgments of how intimidating they found the targets. The accuracy of these

intimidation ratings at predicting the targets' trait aggression was compared for younger adults ($N=87$, $M_{Age}=20.24$ years, $SD_{Age}=2.40$) and older adults ($N=39$, $M_{Age}=71.49$ years, $SD_{Age}=7.59$ years). It was found that there was no meaningful difference between the two age groups, in accuracy in judging trait aggression, with both groups performing highly accurately. It was notable that there was less of a consensus in ratings made by older adults; younger adults agreed more in their ratings. Chapter 5 demonstrated that accuracy improves dramatically in young adults and this chapter demonstrated that this accuracy in detecting trait aggression is not lost with age. The results of this chapter alone demonstrate a positive message for older adults; they can distinguish between more and less aggressive individuals.

Chapter 7: Describing differences in judges and targets during the detection of trait aggression. The final experiment used data to demonstrate the statistical limitations of analysing accuracy of judgment data using nomothetic analyses. 58 participants stood (either facing the presentation screen [$N=29$] or oriented away from the screen and having to look over their shoulder at the target videos [$N=29$]) and watched full videos of 22 walking targets. They judged how threatening and masculine they found the targets. Whilst there was no difference between the accuracy of participants who were facing the target and those who had to look over their shoulder at the target, the results of this study aptly demonstrated the differences in various statistical analyses of judgment data. Analyses of the average threat rating received by the target (as is typically utilized in the literature, known as 'nomothetic' analysis) showed high accuracy in detecting trait aggression and a strong relationship between perceived masculinity and threat of a target. However, analyses of individual judges (known as 'idiographic judge' analysis) showed that some judges had relatively poor accuracy, and some participants displayed notable inaccuracies; judging a more aggressive target as less threatening. Further, individual judge analyses

showed that not all participants adhered to the same ‘masculine is dangerous’ heuristic, with some participants having no strong association between the targets they judged as masculine and those they judged as threatening. A final analysis showed how, when analysing accuracy from all targets, some targets contributed disproportionately to overall accuracy. These targets had extremely ‘salient’ traits (idiographic target analysis) and the removal of just one particularly salient target greatly changed the findings for group accuracy. This chapter provides the most clear analytical advice for future research, overviewing and critiquing the issues with nomothetic analysis (as used in Chapter 3) and idiographic judge analysis (as used in Chapters 4, 5 and 6) as well as highlighting the usefulness of idiographic target analysis (beyond data description.)

Chapter 8: General Discussion. The thesis concludes with an overview of the theoretical value of the research programme as a whole. Specifically, it summarises how the work presented here compliments and expands the extant literature into the accuracy of aggression detection. There is also an emphasis on the applied relevance of the work, with a description of potential applications of the research programme to interpersonal judgments of aggression and to the detection of impending aggressors by remote observation. The general discussion includes commentary on the methodological limitations of the thesis, focusing on the limited realism of the work and the problems inherent with laboratory measures of aggression. In an attempt to address these limitations, general advice for future research is presented, with suggestions for improving the context of judgment experiments, a recommendation to use literature based techniques to better explain the communication of traits and suggestions for research with grouped judges and targets.

References

- Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of personality and social psychology*, 55(3), 387.
- Back, M. D., Schmukle, S. C., & Egloff, B. (2010). Why are narcissists so charming at first sight? Decoding the narcissism–popularity link at zero acquaintance. *Journal of Personality and Social Psychology*, 98(1), 132–145. doi:10.1037/a0016338
- Blanchette, I. (2006). Snakes, spiders, guns, and syringes: How specific are evolutionary constraints on the detection of threatening stimuli? *The Quarterly Journal of Experimental Psychology*, 59(8), 1484–1504. doi:10.1080/02724980543000204
- Boshyan, J., Zebrowitz, L.A., Franklin, R.G., McCormick, C.M., & Carré, J.M. (2013). Age similarities in recognizing threat from faces and diagnostic cues. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 69(5), 710–718, doi:10.1093/geronb/gbt054.
- Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial Structure Is a Reliable Cue of Aggressive Behavior. *Psychological Science*, 20(10), 1194–1198. doi:10.1111/j.1467-9280.2009.02423.x
- Costall, A. (1995). Socializing affordances. *Theory & Psychology*, 5(4), 467–481.
- Costall, A., & Morris, P. (2015). The “textbook Gibson”: The assimilation of dissidence. *History of Psychology*, 18(1), 1–14. doi:10.1037/a0038398
- Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9(5), 353–356. doi:10.3758/bf03337021
- Damasio, A.R. (1994). *Descartes Error: Emotion, Reason and the Human Brain*. Avon, New York.

- Dawson, J., Hillsdon, M., Boller, I., & Foster, C. (2007). Perceived barriers to walking in the neighbourhood environment and change in physical activity levels over 12 months. *British Journal of Sports Medicine*, 41(9), 562–568.
doi:10.1136/bjsm.2006.033340
- DeLoache, J. S., & LoBue, V. (2009). The narrow fellow in the grass: human infants associate snakes and fear. *Developmental Science*, 12(1), 201–207.
doi:10.1111/j.1467-7687.2008.00753.x
- Dunn, B. D., Dalgleish, T., & Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience & Biobehavioral Reviews*, 30(2), 239-271. Foster, S., Knuiman, M., Hooper, P., Christian, H., & Giles-Corti, B. (2014a). Do changes in residents' fear of crime impact their walking? Longitudinal results from reside. *Preventive medicine*, 62, 161-166. doi:10.1016/j.ypmed.2014.02.011
- Foster, S., Villanueva, K., Wood, L., Christian, H., & Giles-Corti, B. (2014a). The impact of parents' fear of strangers and perceptions of informal social control on children's independent mobility. *Health & place*, 26, 60-68.
doi:10.1016/j.healthplace.2013.11.006
- Foster, S., Villanueva, K., Wood, L., Christian, H., & Giles-Corti, B. (2014b). The impact of parents' fear of strangers and perceptions of informal social control on children's independent mobility. *Health & place*, 26, 60-68.
doi:10.1016/j.healthplace.2013.11.006
- Fox, E., Griggs, L., & Mouchlianitis, E. (2007). The detection of fear-relevant stimuli: Are guns noticed as quickly as snakes? *Emotion*, 7(4), 691–696. doi:10.1037/1528-3542.7.4.691

- Funder, D. C. (1980). On seeing ourselves as others see us: Self-other agreement and discrepancy in personality ratings. *Journal of Personality*, 48(4), 473–493.
doi:10.1111/j.1467-6494.1980.tb02380.x
- Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, 21(3), 177-182. doi: 10.1177/0963721412445309
- Funder, D. C. (Ed.). (1999). *Personality judgment: A realistic approach to person perception*. Academic Press.
- Geniole, S. N., Denson, T. F., Dixon, B. J., Carré, J. M., & McCormick, C. M. (2015). Evidence from Meta-Analyses of the Facial Width-to-Height Ratio as an Evolved Cue of Threat. *PLoS ONE*, 10(7), e0132726. doi:10.1371/journal.pone.0132726
- Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Gibson, J. J. (1951). The Perception of the Visual World. *The American Journal of Psychology*, 64(3), 440. doi:10.2307/1419017
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, 62, 451-482.
- Good, J. M. M. (2007). The Affordances for Social Psychology of the Ecological Approach to Social Knowing. *Theory & Psychology*, 17(2), 265–295.
doi:10.1177/0959354307075046
- Gunns, R. E., Johnston, L., & Hudson, S. M. (2002). Psychopathic Traits and Perceptions of Victim Vulnerability. *Journal of Nonverbal Behavior*, 26(3), 129–158.
doi:10.1023/a:1020744915533
- Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology*, 78(1), 81–91. doi:10.1037/0022-3514.78.1.81

- Haselton, M. G., & Buss, D. M. (2009). Error management theory and the evolution of misbeliefs. *Behavioral Brain Science*, 32(06), 522. doi:10.1017/s0140525x09991440
- Hehman, E., Flake, J. K., & Freeman, J. B. (2015). Static and Dynamic Facial Cues Differentially Affect the Consistency of Social Evaluations. *Personality and Social Psychology Bulletin*, 41, 1123-1134. doi:10.1177/0146167215591495
- Hehman, E., Leitner, J. B., & Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology*, 49(4), 747-754. doi:10.1016/j.jesp.2013.02.015
- Hu C, Wang Q, Short LA, Fu G (2012) Speech Spectrum's Correlation with Speakers' Eysenck Personality Traits. *PLoS ONE*, 7(3): e33906. doi: 10.1371/journal.pone.0033906
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14(2), 201–211. doi:10.3758/bf03212378
- Johnson, K. L., & Tassinari, L. G. (2005). Perceiving Sex Directly and Indirectly: Meaning in Motion and Morphology. *Psychological Science*, 16(11), 890–897. doi:10.1111/j.1467-9280.2005.01633.x
- Johnston, L., (2012). It's the Way You Walk: Kinematic Specification of Vulnerability to Attack. In Johnson, K., & Shiffrar, M. (Eds.). *People Watching*. New York: Oxford University Press.
- Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: a social relations analysis. *Psychological Bulletin*, 102, 390.
- Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: acquaintance and the big five. *Psychological Bulletin*, 116, 245.

- Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, 21(6), 575–580.
doi:10.3758/bf03198740
- Kramer, D., Maas, J., Wingen, M., & Kunst, A. E. (2013). Neighbourhood safety and leisure-time physical activity among Dutch adults: a multilevel perspective. *International Journal of Behavioral Nutrition and Physical Activity*, 10.
doi:10.1186/1479-5868-10-11
- Levesque, M. J., & Kenny, D. A. (1993). Accuracy of behavioral predictions at zero acquaintance: A social relations analysis. *Journal of Personality and Social Psychology*, 65, 1178.
- Lipp, O. V., Derakshan, N., Waters, A. M., & Logies, S. (2004). Snakes and Cats in the Flower Bed: Fast Detection Is Not Specific to Pictures of Fear-Relevant Animals. *Emotion*, 4(3), 233–250. doi:10.1037/1528-3542.4.3.233
- McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology*, 43, 293–303.
- McNaughton, N., & Corr, P. J. (2004). A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance. *Neuroscience & Biobehavioral Reviews*, 28(3), 285–305.
- Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, 130(3), 466–478.
doi:10.1037/0096-3445.130.3.466
- Plant, E. A., Goplen, J., & Kunstman, J. W. (2011). Selective Responses to Threat: The Roles of Race and Gender in Decisions to Shoot. *Personality and Social Psychology Bulletin*, 37(9), 1274–1281. doi:10.1177/0146167211408617

- Schaller, M., Faulkner, J., Park, H. J., Neuberg, L. S., & Kenrick, T. D. (2004). Impressions of Danger Influence Impressions of People: An Evolutionary Perspective on Individual and Collective Cognition. *Journal of Cultural and Evolutionary Psychology*, 2(3), 231–247. doi:10.1556/jcep.2.2004.3-4.4
- Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., & Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1656), 575-584.
- Swann, W. B., Jr. (1984). Language and Social Knowing. *PsycCRITIQUES*, 29(3). doi:10.1037/022701
- Thoresen, J. C., Vuong, Q. C., & Atkinson, A. P. (2012). First impressions: Gait cues drive reliable trait judgements. *Cognition*, 124(3), 261–271. doi:10.1016/j.cognition.2012.05.018
- Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, 98, 281.
- Zebrowitz, L. A., & Collins, M. A. (1997). Accurate Social Perception at Zero Acquaintance: The Affordances of a Gibsonian Approach. *Personality and Social Psychology Review*, 1(3), 204–223. doi:10.1207/s15327957pspr0103_2
- Zebrowitz, L.A. & Montepare, J.M. (2005). Appearance DOES matter. *Science*, 308, 1565-1566.

2. Evidence of Big Five and Aggressive Personalities in Gait Biomechanics

Based on the published manuscript ;

Satchell, L., Morris, P., Mills, C., O'Reilly, L., Marshman, P., & Akehurst, L. (2016).

Evidence of Big Five and Aggressive Personalities in Gait Biomechanics. *Journal of Nonverbal Behavior*. doi:10.1007/s10919-016-0240-1

Abstract

Behavioural observation techniques which relate action to personality have long been neglected (Furr & Funder, 2007) and, when employed, often use human judges to code behaviour. In the current study we used an alternative to human coding (biomechanical research techniques) to investigate how personality traits are manifest in gait. We used motion capture technology to record 29 participants walking on a treadmill at their natural speed. We analysed their thorax and pelvis movements, as well as speed of gait. Participants completed personality questionnaires, including a Big Five measure and a trait aggression questionnaire. We found that gait related to several of our personality measures. The magnitude of upper body movement, lower body movement and walking speed, were related to Big Five personality traits and aggression. The way that someone walks can be indicative of their personality. Gait can communicate core aspects of Big Five and aggressive personalities. We know of no other examples of research where gait has been shown to correlate with self-reported measures of personality and suggest that more research should be conducted between largely automatic movement and personality.

2.1 Introduction

“...the sense I have of my own importance is carried in the way I swagger. Indeed, some of the most pervasive features of my attitude to the world and to others is encoded in the way I project myself in public space; whether I am macho, or timid, or eager to please, or calm and unflappable.”

Taylor (1995, p.171)

The philosopher Charles Taylor (1995) suggests that there is a relationship between his ‘attitude’ and the way he walks. However, Taylor’s claim has not been subject to thorough empirical investigation. If Taylor is correct, the ‘automatic’ action of walking could contain information about personality traits. We know of only one other study that has explored this interesting question, and in their investigation Thoresen, Vuong and Atkinson (2012) did not find a relationship between gait and the Big Five personality traits. It is possible that Thoresen et al. did not find a relationship between personality and gait due to their choice of gait analysis. Thoresen et al. adapt Troje’s (2002) technique of reducing whole body motion down to a smaller number of parameters for analysis. This method of analysis does not include the information available from examining relative movements of the body, such as relative upper body to lower body motion. In our experiment, we utilised gait analysis techniques typical of the gait literature (such as; Goujon-Pillet, Sapin, Fodé & Lavaste, 2008) that include this information which could demonstrate the relationship between gait and personality.

Behavioural observation is an established research technique (albeit normally for actions more deliberate or more automatic than walking) in individual differences psychology for judging personality from action. It has been claimed that behavioural observation has been neglected as a research technique in trait psychology experiments for

some time (Furr & Funder, 2007). Behavioural observation techniques require a coder to rate participants' actions and behaviours, such as facial movement (Ekman, Friesen & Hager, 2002) or 'social engagement' (Funder, Furr & Colvin, 2000) for analysis. However, using human coders to recognise the 'relevant' (Funder, 1999) cues to personality has room for error, as some important detail of movement may not be salient or of interest to a human coder. Here, we code the movements of individuals with standard gait biomechanics measures and attempt to demonstrate a relationship between movement and personality.

Our intention is to deliver proof of concept research that demonstrates the relationship between gait and personality. We focused on the relationship between the Big Five (using the measure created by; John, Naumann & Soto, 2008) and gait. We also investigated the relationship between gait and dispositional aggression (using the measure created by; Buss & Perry, 1992). It would be beneficial if there were cues to an approaching stranger's inclination to aggression in their gait. Further, Troscianko et al. (2004) encourage the investigation of any potential relationship between an individual's biological motion and their intention to engage in aggression. Their research finds that 'distinctive gait' is important for predicting impending criminal acts, but Troscianko et al. (2004) do not provide analyses to demonstrate what this distinctive gait may be.

The current research was exploratory with no firm hypotheses. Instead, we explored various aspects of gait and their relationships with aggressive and Big Five personalities. It is reasonable to expect some relationships between sex typical personality and sex typical motion, such as expecting the known differences in trait aggression (with men typically being more aggressive, Eagly & Steffen, 1986; Wilson & Daly, 1997) to relate to the sex typical upper body or 'thorax' movement (with men typically having more upper body rotation within the horizontal plane and women having more pelvis rotation

within the horizontal plane, Bruening, Frimenko, Goodyear, Bowden, & Fullenkamp, 2015). Thus, it is possible, for example, that those in our sample with greater thorax movement (which is seen to be male typical) will report male-typical high trait aggression. However, it is difficult to predict how movement variation *within* the sexes would relate to traits and therefore we form no specific hypotheses.

2.2 Method

2.2.1 Participants. Twenty nine participants were recruited (male = 16, $M_{Age} = 21.14$ years, $SD_{Age} = 2.28$ years) for the experiment. Participants were given a £5 shopping voucher for their participation.

2.2.2 Materials. Participants completed the Buss-Perry Aggression Questionnaire (Buss & Perry, 1992, analysed using revisions suggested by; Bryant & Smith, 2001, see Appendix F) and the Big Five Inventory (John, Naumann & Soto, 2008). The Buss-Perry Aggression Questionnaire quantifies trait aggression into four subscales; tendencies to Physical aggression and Verbal Aggression and dispositional Anger and Hostility towards others. The Buss-Perry Aggression Questionnaire is a commonly used psychometric tool to measure trait aggression has good evidence of external validity (Archer & Webb, 2006; Diamond, 2006; O'Connor, Archer & Wu, 2001). The Big Five Inventory is a measure of the widely used Big Five traits; Conscientiousness, Agreeableness, Neuroticism, Openness to Experience and Extraversion. The Big Five model is the most widely accepted and researched personality model in individual differences research (Funder, 2001).

2.2.3 Procedure. 16 cameras, sampling at 200 Hz, using Qualysis motion capture technology (Qualysis, Sweden), were used to film participants walking for 60 seconds on a treadmill. Reflective markers (used to track the motion of the upper body (thorax) and lower body (pelvis)) were placed on the participants. These included the markers located on the thorax and pelvis, (figure 3.1) and markers on the feet to track gait cycles.

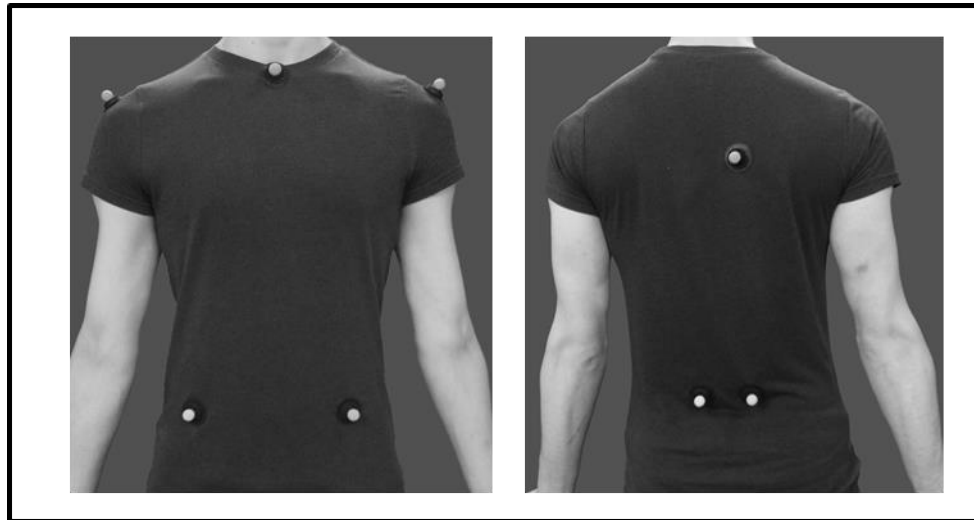


Figure 2.1. Placement of retroreflective markers. LEFT, the anterior view of marker placements, showing retroreflective markers placed on the acromion process of both shoulders (thorax reference point), the suprasternal notch (thorax reference point) and the crests of the anterior iliac spines (pelvis reference point). RIGHT, the posterior view of marker placements, showing retroreflective markers placed on the posterior crests of the iliac spines (pelvis reference point) and a tracking marker placed right of the spine on the upper back (overall reference point).

Participants were asked to adjust the treadmill speed to that of their normal gait (Gait Speed in km/h) and were given time to familiarise themselves with walking on a treadmill. When participants felt ready, the motion capture equipment recorded their walk for 60 seconds. Subsequently, participants completed the Buss-Perry Aggression Questionnaire and the Big Five Inventory and were debriefed.

2.2.4 Analyses.

2.2.4.1 Statistical analyses. As aspects of movement may not be normally distributed we opted to use nonparametric analyses for our results; Mann-Whitney U tests for the demonstration of sex differences and Spearman's ranked correlations to

demonstrate the relationship between psychometrics and gait. Nonparametric tests are better suited to analysing smaller sample sizes.

2.2.4.2 Gait analyses. From the 60 seconds of walking, the first five uninterrupted gait cycles (as defined by change in the velocity of heel markers, see Zeni, Richards & Higginson, 2008) were used for analysis. To describe key gait characteristics associated with the Big Five and trait aggression, we calculated the range of motion (ROM) of the thorax relative to the fixed horizontal plane, the laboratory (Thorax-Lab ROM, see figure 2.2), to give a measure of upper body movement. We also calculated the range of motion of the pelvis relative to the fixed horizontal plane (Pelvis-Lab ROM, see figure 2.2) for a measure of lower body movement. Further, we calculated the range of motion of the thorax (upper body) relative to the pelvis (lower body) in the horizontal plane (Thorax-Pelvis ROM, see figure 2.2). The Thorax-Pelvis ROM is calculated as the difference between the angle of Thorax and Pelvis ROM when the Thorax is at its greatest during a gait cycle. Figure 2.3 presents the content of Figure 2.2 in a way that represents this measure from an analytical perspective.

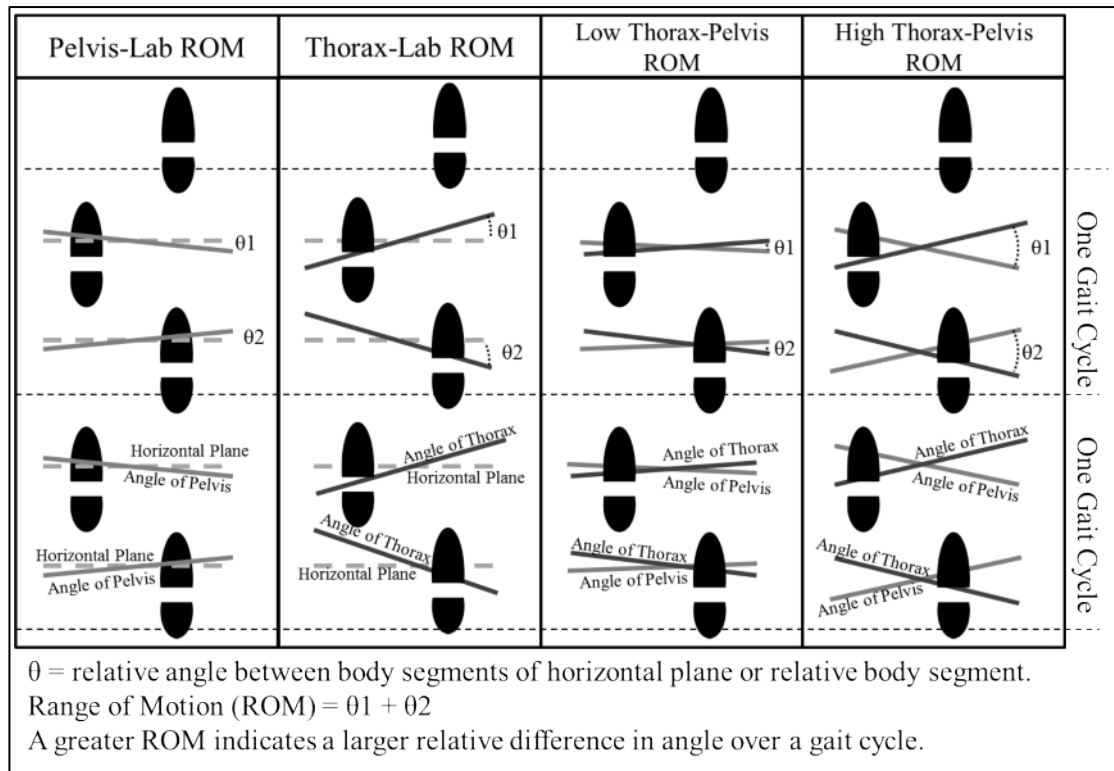


Figure 2.2. A graphical representation of the Range of Motion measurements used

in this experiment.

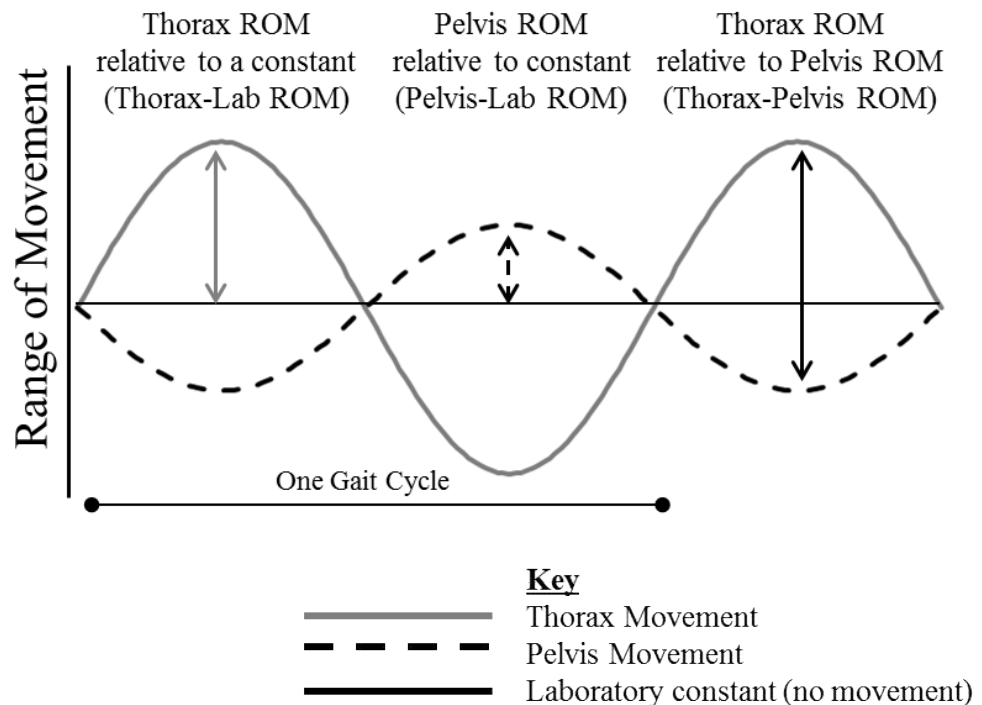


Figure 2.3. An alternative representation to that in Figure 2.2, presenting the Range

of Motion measurements use in this study.

From each gait cycle, ranges of motion (ROM) were calculated by finding the difference between the minimum and maximum angle of movement during each gait cycle. The ROM was averaged across the five gait cycles to produce a single value for analysis per participant. These measures are typical of research in gait biomechanics (see similar analyses by Goujon-Pillet, Sapin, Fodé & Lavaste, 2008).

2.3. Results

2.3.1 Sample Sex Differences in Personality and Gait. We found some notable sex differences in the biomechanical (such as, Thorax-Pelvis ROM) and psychometric (such as, Neuroticism and Hostility) measures (see table 3.1) of our study, but few of these were moderate or large effects (see; Ferguson, 2009). Due to the evidence of sex differences from our own sample, as well as of the known sex differences in gait style (Bruening, Frimenko, Goodyear, Bowden, & Fullenkamp, 2015; Troje, 2002) and physical aggression (Eagly & Steffen, 1986; Wilson & Daly, 1997) in the existing literature, we chose to analyse the data from females and males separately in addition to analysing the pooled data

2.3.2 Aggression and Motion Correlations. We tested for Spearman's ranked correlations between biomechanical aspects of gait and trait aggression. The results showed that participants Thorax-Pelvis ROM reflected the trait physical aggression of individuals, especially with female participants (see table 3.2). It should be noted that it was only the relative motion of thorax and pelvis that correlated with aggression; neither thorax movement on its own nor pelvis movement on its own correlated with aggression. Rather, the increased relative movement of the thorax (upper body) and pelvis (lower body) together was reflective of increased physical aggression. There was a moderate

correlation between gait speed and aggression for males but no correlation between gait speed and aggression for females or the sample as a whole.

Table 2.1. *The Mann Whitney U Differences Between Sexes for the Biomechanical and Psychometric Measures*

Variable	U	r	Female Participants		Male Participants	
			Mean	SD	Mean	SD
Thorax-Pelvis ROM (°)	58.00	.37	7.71	2.94	5.63	2.06
Thorax-Lab ROM (°)	77.00	.22	8.28	2.34	7.35	2.16
Pelvis-Lab ROM (°)	68.00	.28	7.22	1.89	5.95	1.92
Gait Speed (km/h)	98.50	.04	3.40	0.73	3.44	0.71
Physical Aggression	74.50	.23	24.69	7.62	28.81	9.17
Verbal Aggression	60.50	.33	14.69	4.48	18.38	5.57
Anger	85.00	.14	22.15	4.06	21.44	6.67
Hostility	52.00	.38	29.31	8.89	21.31	8.91
Conscientiousness	93.50	.08	31.54	5.63	32.19	4.53
Agreeableness	76.00	.20	35.00	4.65	33.56	3.31
Neuroticism	42.50	.43	27.77	4.11	19.44	8.92
Openness	79.50	.17	34.15	4.34	32.13	5.08
Extraversion	95.50	.06	25.85	4.04	25.31	6.76

Note: Ranges of motion are as follows; Thorax-Pelvis is the range of Thorax motion relative to the participants' pelvic motion, Thorax-Lab is the range of Thorax motion relative to the horizontal plane (laboratory) and Pelvis-Lab range is the range of pelvic motion relative to the horizontal plane (laboratory)

2.3.3 Big Five Factors and Motion Correlations. There were some elements of gait that were powerfully correlated with the Big Five traits (using Spearman's ranked correlations, see table 3.2). Notably large correlations included, the strong positive correlations between agreeableness and Pelvis-Lab ROM (.80) and conscientiousness and Thorax-Lab ROM (-.69) for female participants, and extraversion and Thorax-Lab ROM (.58) for male participants.

Table 2.2. *Spearman's Ranked Correlations Between Biomechanics and Psychometrics Across Participants, Split by Gender and Together as an Overall Sample*

	Ranges of Motion ¹			
Trait	Thorax-Pelvis	Thorax-Lab	Pelvis-Lab	Gait Speed
Male Participants (N=16)				
Physical Aggression	0.47	0.39	0.26	0.33
Verbal Aggression	0.42	<u>0.72</u>	<u>0.67</u>	-0.47
Anger	0.34	0.35	0.43	0.35
Hostility	-0.23	0.17	0.31	-0.37
Conscientiousness	-0.31	-0.18	-0.12	0.07
Agreeableness	0.05	-0.01	0.01	-0.02
Neuroticism	-0.23	-0.23	-0.09	-0.21
Openness	-0.23	-0.42	-0.29	0.17
Extraversion	0.11	<u>0.58</u>	0.49	0.11
Female Participants (N=13)				
Physical Aggression	<u>0.74</u>	0.03	-0.20	-0.07
Verbal Aggression	0.08	-0.15	-0.13	-0.05
Anger	<u>0.55</u>	0.06	-0.12	0.17
Hostility	-0.02	-0.42	-0.47	-0.22
Conscientiousness	-0.21	<u>-0.69</u>	0.14	<u>0.55</u>
Agreeableness	0.06	-0.47	<u>0.80</u>	<u>0.64</u>
Neuroticism	0.01	0.33	-0.28	-0.44
Openness	0.23	-0.43	0.02	-0.03
Extraversion	0.26	-0.37	<u>0.54</u>	<u>0.51</u>
Overall Sample (N=29)				
Physical Aggression	0.43	0.23	0.02	0.12
Verbal Aggression	0.08	0.20	0.18	-0.19
Anger	0.48	0.26	0.21	0.27
Hostility	0.09	0.02	0.07	-0.31
Conscientiousness	-0.21	<u>-0.52</u>	0.06	0.39
Agreeableness	0.06	-0.20	0.48	0.29
Neuroticism	0.01	0.05	0.01	-0.27
Openness	0.03	-0.28	-0.07	0.03
Extraversion	0.13	0.18	0.43	0.23

¹ Ranges of motion are as follows; Thorax-Pelvis range is the range of Thorax movement relative to the participants' range of pelvic movement, Thorax-Lab range is the range of Thorax movement relative to the lab and Pelvis-Lab range of motion is the range of pelvic movement relative to the lab.

Bold type highlights effects equal to larger than Ferguson's (2009) minimum recommended effects and **Bold and Underlined** type highlights effect equal to or larger than Ferguson's moderate effects.

2.4 Discussion

Taylor (1995) claimed that the “swagger” (p.171) of his walk reflected his personality. In our biomechanical exploration of gait, we provide empirical evidence that personality is indeed manifest in how a person walks. Aspects of gait, such as (i) relative movement between the upper and lower body (Thorax-Pelvis ROM), (ii) upper body movement alone (Thorax-Lab ROM), (iii) lower body movement alone (Pelvis-Lab ROM), and (iv) gait speed, related to differing aspects of our participants’ personality. Importantly, our use of biomechanics allowed a measurement approach to analysing gait that was not reliant on human coders. However, this technology does record more information than a human could process (our equipment captures movement at 200 Hz), suggesting that our results may present the available cues to personality in biomechanical recording and not the available cues to personality in visual perception. This concern was also raised by Thoresen, Vuong and Atkinson (2012) which is why they chose to analyse their movement based on what was visually salient. In our study, we opted to use the gait analysis more typical of the biomechanics literature as we are interested in the relationship between the actual movement involved in gait and personality (rather than how people make judgements about gait), which may explain the difference between our findings and those of Thoresen et al. We also examined different aspects of movement, such as our Thorax-Pelvis ROM, instead of Thoresen et al.’s compound whole body measures, which may have allowed us further sensitivities to gait mechanics relevant to our measured traits.

The walks of participants who self-reported high physical aggression comprised greater relative movement between the upper and lower body (higher Thorax-Pelvis ROM). Importantly, this relationship was not simply a property of heightened shoulder movement (Thorax-Lab ROM) or pelvis movement (Pelvis-Lab ROM) as these variables alone were not predictive of aggression; it was the relative motion of the upper and lower

body that was important. When walking, the body naturally rotates a little; as an individual steps forward with their left foot, the left side of the pelvis will move forward with the leg and the left shoulder will move back and the right shoulder forward to maintain balance. Put simply, an aggressive walk is one where this rotation is exaggerated. These findings support Troscianko et al.'s (2004) suggestion that there are aspects of biological motion that could be useful to detect in crime prevention. If the CCTV observers in their experiments could be trained to recognise the aggression-relevant gait demonstrated in this research, their ability to recognise impending crimes (which is already high) could be improved further.

It might be assumed that traits that have more associations with activity (e.g. extraversion) have stronger associations with gait than those that do not (e.g. agreeableness or verbal aggression). It was therefore a surprise to find such strong correlations between agreeableness and gait in women, and verbal aggression and gait in men. Given our rather small (but sufficient) sample size future research could attempt to replicate these relationships as the robust, strong correlations between these variables would be noteworthy for personality psychology, movement psychology and even in real world settings (such as with Troscianko et al.'s CCTV research).

It is plausible that gait affects personality through the embodiment of a walking style, for example, adopting a confident style of gait and then self-rating high extraversion (see similar work on 'Power poses'; Cesario & McDonald, 2013). Further research is required to establish whether gait affect personality or personality affects gait. It could also be the case that gait affects how participants complete self-report measures, with feelings of aggression or confidence (neuroticism or extraversion) being ameliorated or diminished by recently having their gait observed.

Our findings of correlations between gait and personality are interesting, as gait is a behaviour that holds an unusual place between the automatic, physiological correlates of personality (such as, cerebral blood flow; Johnson et al., 2014) and the behavioural or social engagement correlates of personality (such as, ‘acting playfully’; Furr & Funder, 2012). To most people walking is a relatively automatic behaviour yet it is reflective of individual psychology, such as one’s openness to new experiences. Our findings are important in explaining accurate interpersonal trait judgments and our objective methods could be used with other potentially personality-relevant movements, such as seated pose and shaking hands. The existence of these correlations demonstrates the potential of research into relationships between individual differences in psychology and individual differences in movement.

2.4.1 Conclusion.

In this exploratory paper, we found a large number of strong and significant correlations between various aspects of gait and pen and paper psychometric measures. This suggests that walking style and personality are related. We avoid extensive theoretical interpretation of the particular relationships between aspects of gait and personality traits revealed in the study, instead our conclusion is a broad one; personality is manifest in the way we walk.

References

- Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of personality and social psychology*, 55(3), 387.
- Albright, L., Malloy, T. E., Dong, Q., Kenny, D. A., Fang, X., Winkquist, L., & Yu, D. (1997). Cross-cultural consensus in personality judgments. *Journal of personality and social psychology*, 72(3), 558.
- Archer, J., & Webb, I. A. (2006). The relation between scores on the Buss–Perry Aggression Questionnaire and aggressive acts, impulsiveness, competitiveness, dominance, and sexual jealousy. *Aggressive Behavior*, 32, 464–473.
doi:10.1002/ab.20146
- Biesanz, J. C., West, S. G., & Millevoi, A. (2007). What do you learn about someone over time? The relationship between length of acquaintance and consensus and self-other agreement in judgments of personality. *Journal of personality and social psychology*, 92(1), 119.
- Bruening, D. A., Frimenko, R. E., Goodyear, C. D., Bowden, D. R., & Fullenkamp, A. M. (2015). Sex differences in whole body gait kinematics at preferred speeds. *Gait & Posture*, 41(2), 540–545. doi:10.1016/j.gaitpost.2014.12.011
- Bryant, F. B., & Smith, B. D. (2001). Refining the architecture of aggression: a measurement model for the Buss–Perry Aggression Questionnaire. *Journal of Research in Personality*, 35(2), 138-167.
- Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of personality and social psychology*, 63(3), 452.
- Cesario, J., & McDonald, M. M. (2013). Bodies in context: Power poses as a computation of action possibility. *Social Cognition*, 31(2), 260-274.

- DeGroot, T., & Gooty, J. (2009). Can nonverbal cues be used to make meaningful personality attributions in employment interviews?. *Journal of Business and Psychology*, 24(2), 179-192.
- Diamond, P. M. (2006). The Short-Form Buss-Perry Aggression Questionnaire (BPAQ-SF): A Validation Study With Federal Offenders. *Assessment*, 13, 227–240.
doi:10.1177/1073191106287666
- Eagly, A. H., & Steffen, V. J. (1986). Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin*, 100(3), 309–330. doi:10.1037/0033-2909.100.3.309
- Eagly, A. H., & Steffen, V. J. (1986). Gender and aggressive behavior: a meta-analytic review of the social psychological literature. *Psychological bulletin*, 100(3), 309.
- Ekman, P., Friesen, W. V., & Hager, J. C. (2002). FACS investigator's guide. *A human face*.
- Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), 532.
- Funder, D. C. (Ed.). (1999). *Personality judgment: A realistic approach to person perception*. Academic Press.
- Funder, D. C. (2001). Personality. *Annual Review of Psychology*, 52, 197-221.
- Funder, D. C., Furr, R. M., & Colvin, C. R. (2000). The Riverside Behavioral Q sort: A Tool for the Description of Social Behavior. *Journal of Personality*, 68(3), 451-489.
- Furr, R.M., & Funder, D.C. (2007). Behavioral Observation In R.W. Robins, R.C. Fraley & R.F. Krueger (Ed.) *Handbook of Research Methods in Personality Psychology* (pp. 273-291) New York , NY: The Guilford Press.
- Goujon-Pillet, H., Sapin, E., Fodé, P., & Lavaste, F. (2008). Three-Dimensional Motions of Trunk and Pelvis During Transfemoral Amputee Gait. *Archives of Physical Medicine and Rehabilitation*, 89(1), 87–94. doi:10.1016/j.apmr.2007.08.136

- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114-158). New York, NY: Guilford Press.
- Johnson, D. L., Wiebe, J. S., Gold, S. M., Andreasen, N. C., Hichwa, R. D., Watkins, G. L., & Ponto, L. L. B. (2014). Cerebral blood flow and personality: a positron emission tomography study. *The American Journal of Psychiatry*, 156(2), 252-7
- McNaughton, N., & Corr, P. J. (2004). A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance. *Neuroscience & Biobehavioral Reviews*, 28(3), 285–305.
- O'Connor, D. B., Archer, J., & Wu, F. W. C. (2001). Measuring aggression: Self-reports, partner reports, and responses to provoking scenarios. *Aggressive Behavior*, 27, 79–101. doi:10.1002/ab.2
- Taylor, C. (1995). *Philosophical arguments*. Harvard University Press.
- Thoresen, J. C., Vuong, Q. C., & Atkinson, A. P. (2012). First impressions: Gait cues drive reliable trait judgements. *Cognition*, 124(3), 261–271. doi:10.1016/j.cognition.2012.05.018
- Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, 2(5), 2–2. doi:10.1167/2.5.2
- Troscianko, T., Holmes, A., Stillman, J., Mirmehdi, M., Wright, D., & Wilson, A. (2004). What happens next? The predictability of natural behaviour viewed through CCTV cameras. *Perception*, 33(1), 87–101. doi:10.1068/p3402
- Wilson, M., & Daly, M. (1997). Life expectancy, economic inequality, homicide, and reproductive timing in Chicago neighbourhoods. *British Medical Journal*, 314, 1271-1274.

Zeni, J. A., Richards, J. G., & Higginson, J. S. (2008). Two simple methods for determining gait events during treadmill and overground walking using kinematic data. *Gait & posture*, 27(4), 710-714.

3. The Accuracy of Interpersonal Threat Judgments for Detecting Trait Aggression

Manuscript in preparation for publication

Satchell, L., Morris, P., Morrison, E., & Akehurst, L. (in preparation) *The accuracy of interpersonal threat judgments for detecting trait aggression.*

Abstract

Decisions about the risk posed by an approaching stranger should be made quickly and efficiently. The first study in this paper investigated the accuracy of interpersonal threat judgments at reflecting the trait aggression of an approaching person. The second study explores whether physical appearance or gait style best facilitate accurate judgments of trait aggression. In the first study, participants were asked to judge the threat posed by targets who had been filmed walking towards the camera. We also correlated the threat judgments with the self-reported trait aggression scores of the targets to test for accuracy of threat judgments. In the second study some participants were presented with only static information (still images of the targets) and some were presented with limited dynamic information (point light videos of the targets). In the first study threat judgments accurately reflected the trait aggression of the targets. In the second study the judgments of threat made by participants who observed only the dynamic information did reflect targets' trait aggression, but not the judgments made by those in the static information. Judgments of the threat posed by a target can reflect the trait aggression of an approaching person. The results of these two studies demonstrate that judgments of threat are most accurate when gait information is available.

3.1 Introduction

When walking down a street, an approaching stranger could be seen as a potential threat to personal safety. An incorrect judgment about such a threat could lead to injury or worse. The decision about the potential threat a person poses has to be made accurately but also efficiently. As another person approaches, the time to make and act on a judgment becomes increasingly limited and in situations of anxiety the time to make a judgment could be perceived to be even more limited (see ‘defensive distance’; McNaughton & Corr, 2004). When decision making is constrained by time, an individual may use decision making shortcuts or ‘heuristics’ to reach a fast but potentially inaccurate choice (Gigerenzer & Gaissmaier, 2011). Therefore, this paper presents two studies to explore if judgments about the threat another person poses can be accurate, and if accuracy is more influenced by body image (such as appearance or biometric masculinity) or by motion.

Previous research into perceptions of aggression (Carré, McCormick & Mondloch, 2009) and intimidation (Hehman, Leitner & Gaertner, 2013) from faces would suggest that the perceived danger posed by a stranger could be driven by the heuristic that ‘masculine is dangerous’. That is to say, a more masculine-typical body shape (or, in the above cases, the shape of the face) leads to judging a target as more aggressive or more intimidating. This is a legitimate, albeit overly broad, heuristic perhaps driven by the asymmetry in danger posed by the sexes (Archer, 1991). There is evidence for this difference in aggression tendency in research settings, with men being more aggressive in laboratory studies (Buss & Perry, 1992; Eagly & Steffen, 1986). Further, crime statistics demonstrate that men are more likely than women to commit both violent and non-violent crime (Daly & Wilson, 1990; Wilson & Daly, 1985; 1993). Previous research has established that a ‘masculine is dangerous’ heuristic, effects judgments of intimidation and aggression (Carré, McCormick & Mondloch, 2009; Hehman, Flake & Freeman, 2015; Hehman, Leitner & Gaertner,

2013). As many men do not commit aggressive acts, judging all men as threatening and avoiding all unknown, approaching men is a waste of a pedestrian's time and energy. However, the assumption that men are more dangerous than women could support or encourage accurate judgments of threat. It is therefore of interest to explore the influence of the 'masculine is dangerous' heuristic on interpersonal threat judgments.

The previous research on judgments of the malevolent traits of others is entirely based on judgments of intimidation or aggression from photographs of faces (e.g. Carré, et al., 2009; Hehman, et al., 2015; Hehman, et al., 2013). More than this, some studies use extremely brief presentations of faces (39 microseconds; Carré et al., 2009). This choice of stimuli misrepresents the quality of information available from other people in everyday life. An approaching person would offer full body, movement information on their approach (which would likely last longer than 39 microseconds). Part of the novelty of our research is our utilisation of stimulus people who are filmed walking towards camera, from a distance. It has been noted that there has been "relatively little attention in mainstream social psychology to the question of just how informative kinematics (body movements in general) are to the perceiver of human social actions" (Good, 2007, p.267). The walk of an unknown person has been considered important in research on perceptions of vulnerability (arguably the directly opposite phenomenon to threat). Gunns, Johnston and Hudson (2002) show that vulnerability can be reliably judged at a distance, and they report that vulnerability judgments are associated with the femininity of the target's gait. Here, we present the targets of threat judgments as *approaching* strangers in order to understand the influence of the 'masculine is dangerous' heuristic on threat judgments when targets are observed from a distance. Further, we explore whether accurate judgments of aggression (such as seen in previous face photograph research; Carré et al., 2009) can be made from a distance.

We define an accurate threat judgment made by a judge as a judgment which relates to the targets' self-reported trait aggression. This definition of accuracy is typical of the personality judgment literature, where accuracy is defined as the relationship between a self-reported trait of a target and a judge's rating of that target (Funder, 2012). Our measure of trait aggression is the widely used Buss-Perry Aggression Questionnaire (Buss & Perry, 1992), a self-report measure which has been shown to relate to hypothetical (Archer & Webb, 2006; O'Connor, Archer & Wu, 2001) and historic (Diamond, 2006) aggression and is frequently used in contemporary research to report trait aggression (such as; Greitemeyer, 2015; Lake, Stanford, & Patton, 2014, Waldron, Scarpa, Lorenzi, & White, 2015; Zajenkowski & Zajenkowski, 2015).

Much of the existing research into judgments of intimidation or aggression does not attempt to quantify the accuracy of participants' judgments (*c.f.* Carré et al., 2009). Measuring accuracy would allow researchers to investigate whether the 'masculine is dangerous' heuristic is accurate in itself or too broad and a detraction from potentially accurate judgments. In the personality psychology literature, there is research that demonstrates that individuals can make accurate judgments about the traits of another person with limited interaction and information (Cheek, 1982; Funder, 1980; McCrae, 1982; Vazire & Mehl, 2008). Previous research has explored accuracy in judging the Big Five personality traits of walking people presented as point light displays (Thoresen, Vuong & Atkinson, 2012). Whilst Thoresen et al. found that participants were inaccurate at recognising the traits of a target, they also found that participants were highly reliable in their judgments of the targets. This would suggest that, as a sample, participants interpreted the gait they observed in the same way but their interpretation was incorrect. Thoresen et al.'s results suggest that Big Five personality traits cannot be recognised from gait, but it could be the case that participants could recognise trait aggression through judgments of

threat. Given that this thesis has demonstrated that there are aspects of gait that correlate with trait aggression (Chapter 2), it is expected that judgments of threat based on gait information would be accurate. This expectation is supported by research that has suggested that ‘distinctive gait’ is an important feature in accurately detecting impending crimes when watching CCTV footage (Troscianko et al., 2004). Perhaps individuals could find it easier to recognise trait aggression than Big Five personality because they have more experience in trying to identify dispositional aggression in others. In everyday life it is more routine and important to judge the danger an approaching person may pose than recognising the extraversion or openness to experience of a stranger.

Funder (1999) explains how judgments of another’s trait can be accurate with the Realistic Accuracy Model (RAM). The RAM emphasises the importance of the target person in interpersonal judgments. The RAM suggests that a judgment of a target’s traits can be made if the target is revealing *relevant* information to the judge. This information is *available* to and *detected* by a perceiver and is then *utilized* to form an accurate judgment. This approach to understanding accuracy of trait judgment reinforces the potential consequences of limiting the *available* information for a judge. It is difficult to conclude that face width is important for judges to utilize in aggression judgments (e.g. Carré et al., 2009), when experiments limit the information available to the judges as only the face of the targets. Therefore, in the second experiment presented in this paper we explore the influences of limiting the availability of information from another person. We do this by explore the accuracy of threat judgments when judges are only presented with body image or movement information.

Previous research into judgments of threat does not use realistic presentations of a potential aggressor. Here, we present two experiments to improve everyday relevance of the paradigms widely used in the threat judgment literature. In the first study we explore

how judgments of threat are influenced by a ‘masculine is dangerous’ heuristic when the judgments are based on observing an approaching person. This study also reports the accuracy of the threat judgments at predicting targets’ trait aggression. In our second study, we further explore the importance of the *available* information from targets to judges when judging threat. We do this by limiting targets to observing solely static or solely dynamic information. The objective of this second experiment is to demonstrate if appearance or movement is more important to detecting trait aggression in threat ratings.

3.2 Study One: The ‘Masculine is Dangerous’ Heuristic and the Accuracy of Threat Judgments

Firstly, to explore the ‘masculine is dangerous’ heuristic we investigated the relationships between (i) targets’ body shape and targets’ self-reported trait aggression and (ii) targets’ body shape and participants’ threat ratings. Secondly, we investigated the accuracy of our participants’ threat judgments by comparing their judgments to the trait aggression scores of the targets.

3.2.1 Method

3.2.1.1 Participants. 61 undergraduate student participants (female= 47, $M_{Age}=19.18$ years, $SD_{Age}=3.34$ years) were recruited from a participant pool at a UK university. They were compensated for their participation with course credit.

3.2.1.2 Design. This study employed a correlational design. We correlated targets’ body shape (see 3.2.1.3 for the body shape index) and targets’ self-reported aggression. We also correlated our index of how masculine typical the targets’ bodies were with the average threat rating that each target received. Finally, we tested the accuracy of the threat judgments by correlating the average threat rating received by each target with targets’ self-reported aggression.

3.2.1.3 Materials. We used the same 23 (female= 12, $M_{Age}= 20.57$ years, $SD_{Age}= 2.02$ years) target stimuli in both of our experiments. All targets self-reported being “White” or “Caucasian”. These targets were deliberately chosen to be between the ages of 18 and 25 years to be in an age group likely to commit an aggressive crime and potentially pose a danger (Daly & Wilson, 2001).

We filmed the stimulus targets walking on a treadmill wearing standardised clothes (for male targets a white t shirt and blue shorts, see figure 3.1A, and for female targets a grey vest top and black leggings). All videos were 10 seconds of uninterrupted walking with the targets displaying a neutral facial expression and walking toward the camera (all targets were also wearing retro-reflective markers which were visible in the videos, see figure 1A).

Targets completed the Buss-Perry Aggression Questionnaire (Buss & Perry, 1992, analysed using revisions suggested by Bryant & Smith, 2001, see Appendix F). This inventory was used as a simple but established measure of targets’ aggressive tendencies. From this questionnaire the physical aggression subscale was extracted, as it is most relevant to detecting genuine inter-human threat. As per Bryant and Smith’s (2001) revisions participants could score between 3 and 21 for the physical aggression measure. In our sample of targets there was a reasonable spread of aggression scores ($M_{Aggression}=7.52$, $SD_{Aggression}=4.67$, $Min_{Aggression}=3$, $Max_{Aggression}=19$).

The targets’ height, shoulder diameter, hip diameter and waist diameter were measured and used to compute literature-typical biometric values (Gray & Wolfe, 1980; Hughes & Gallup, 2003). These biometric measures are known to differ between the sexes, with males typically having larger shoulder-to-hip ratios and waist-to hip-ratios (Hughes & Gallup, 2003) and being taller (Gray & Wolfe, 1980) than women. We used this to create a Sexual Dimorphism Index (SDI) score for each target, a measure of how masculine-typical

to feminine-typical our targets' bodies were relative to the other targets in the experiment. This SDI score was derived by first ranking the targets' biometric features, so that the larger the measure to shoulder-to-hip ratio, the waist-to-hip ratio and height, the larger ranked value. These rankings were summed to produce one total ranking measure and then ranked again. This gave a ranking between 1 and 23 for each target, with a higher ranking value implying a more masculine typical morphology. Our metric divided our sample as would be expected; with all targets with an SDI less than 13 all being female and those with an SDI equal to or greater than 13 being male.

3.2.1.4 Procedure. Participants gave written informed consent. Participants made ratings of the 23 full video presentations of the targets on 1-9 Likert scales including the critical judgment of *Threatening-Unthreatening*. The order of target presentation was randomised for each participant.

3.2.2 Results

3.2.2.1 Target Body Shape and Aggression. We found evidence that the 'masculine is dangerous' could heuristic benefit judgments of the male targets with the male targets with more masculine bodies self-reporting higher trait aggression scores ($r(11)=.74$, 95% CI [.15, .96], $p=.009$). However the trait aggression of the female targets ($r(12)=.09$, 95% CI [-.49, .77], $p=.774$) and the targets in general ($r(23)=.19$, 95% CI [-.40, .67], $p=.392$) did not relate to their SDI. There was no difference in the trait aggression between the male ($M_{Aggression}=7.54$, $SD_{Aggression}=4.68$) and female ($M_{Aggression}=7.50$, $SD_{Aggression}=4.89$) targets ($t(21)=.02$, $p=.982$, $d=.01$).

3.2.2.2 Influence of 'masculine is dangerous' heuristic. Each target received similar threat ratings from each of the participants ($\alpha=.93$). It is also important to note that there was significant variation in how participants judged the 23 targets they saw ($F(22,1320)=13.90$, $p < .001$). Overall we found that our participants were judging the more

targets with more masculine typical bodies as more threatening, with a moderate to large size of correlation ($r(23)=.68$, 95% CI [.37, .87], $p<.001$). Further analyses showed that the SDI measure did not significantly correlate to the threat ratings for the male targets ($r(11)=.59$, 95% CI [-.24, .89], $p=.057$), and the female targets ($r(12)=.45$, 95% CI [-.31, .88], $p=.147$) when analysed as separately, but these were still positive correlations of a moderate size.

3.2.2.3 Accuracy at detecting trait aggression. Participants were able to discriminate between more and less aggressive targets. The threat ratings received by the targets correlated with self-reported trait aggression scores, ($r(23)=.43$, 95% CI [.00, .70], $p=.042$) demonstrating some accuracy in threat judgments. The magnitude of the correlation between threat judgements and self-reported aggression were similar for both male targets ($r(11)=.61$, 95% CI [-.05, .90], $p=.047$) and female targets ($r(12)=.41$, 95% CI [-.11, .80], $p=.190$).

3.2.3 Study One Summary Discussion

We found that the threat judgments received by our targets related to the masculinity of their bodies (as measured by the SDI); more masculine targets were judged as more threatening. We also found that the average threat rating received by each target reflected target aggression, demonstrating that participants' threat judgments were accurate at reflecting potential aggressors. What is notable from this finding is that, even in a non-offending population where variance in aggression was not large, participants were able to accurately discriminate between those who were more or less aggressive. Our participants recognised subtle but important differences between target stimuli. Whilst the male targets' bodies may have communicated some of this information, the detection of aggression in female targets without the assistance of the SDI does suggest that there are other factors, beyond body shape, that communicated aggression. With accuracy possible in threat

ratings, it was of interest to investigate what *available* (see RAM) information from the targets was detected by participants to enable the relationship between threat judgments and trait aggression.

3.3 Study Two: Controlling the Availability of Target Information

In our second experiment we isolated static and dynamic target information from the stimuli used in the first experiment, to investigate which was more important on accuracy of threat judgments. In the parlance of the RAM, we were effectively controlling the *available* target information to our judges, which in turn affected what judges could *detect* and *utilize* to form their threat judgments. We separated the target information into static information, (physical appearance and body shape), and dynamic information, (gait style). Participants in the ‘static information alone’ condition saw a single frame (taken from the beginning of that target’s first ‘gait cycle’, their first filmed step forward) from the videos used in study one. The footage for the ‘dynamic information alone’ condition was extracted using motion capture technology to produce point light walkers. By separating out the sources of information we intended to investigate whether the availability of information from the appearance of our targets (static information) or the gait style of our targets (dynamic information) affected the accuracy of threat judgments.

3.3.1 Method

3.3.1.1 Participants. For the second study 40 participants were recruited (female=31, M_{Age} =19.15 years, SD_{Age} = 1.67 years) from a participant pool at a UK university. They were compensated for their participation with course credit. Participants were randomly allocated to one of the two conditions.

3.3.1.2 Design. A one-way between subjects design was used. The differences between the norm participant threat ratings in the two presentation conditions (static information only and dynamic information only) were tested using an ANOVA. There

were also correlational elements of the design which investigated the relationship between the average threat rating received by each target in each condition and the targets' self-reported aggression.

3.3.1.3 Materials. The targets in the second experiment were the same as those in the first. In fact, whilst filming the targets using standard filming, their motion was also captured using infrared filming techniques. This allowed us to have a point light walker for the dynamic information alone condition which was the exact same section of gait used in first experiment. We were also able to then take a frame from the videos used in the first experiment present for 10 seconds as an image for the static information alone condition. This frame was taken from the beginning of the target's first gait cycle.

For the dynamic information alone condition, targets' movement was captured using Qualysis motion capture filming technology. Ten ProReflex infrared cameras captured the targets' movement at 100 Hz. 13 Retroreflective markers were placed on the models' bodies at the lateral malleolus of both ankles, the lateral epicondyle of both knees, both greater trochanters, the acromion processes, both wrists, both elbows and the forehead. Thus, there were three possible conditions of presentation; a video of the target walking along for the full information condition (see figure 3.1A), a frame from the beginning of the target's second gait-cycle for the static information alone condition (see also figure 3.1A), and a point light walker produced from the motion capture filming for the dynamic information alone condition (see figure 3.1B).

3.3.1.4 Procedure. Participants gave written informed consent. Participants were presented with the same 23 walkers as in study one and were, again, asked to rate the targets on 1-9 Likert scales including the critical judgment of Threatening-Unthreatening. The order of target presentation was randomised for each participant. The targets were presented to judges in one of two different formats depending on condition allocation;

‘static information alone’ participants ($N = 20$) saw a single frame from the walking videos for 10 seconds and ‘dynamic information alone’ participants ($N = 20$) saw point light representations of the targets walking for 10 seconds.

3.3.2 Results.

3.3.2.1 How threat ratings varied across conditions. Threat ratings received by the targets were highly similar in the static alone ($\alpha=.90$) and dynamic alone ($\alpha=.83$) conditions. The average threat rating received by the targets in the dynamic alone and static alone conditions did not significantly correlate and showed only a small to moderate effect ($r(23)=.34$, 95% CI [-.03, .66], $p=.111$). Threat ratings in the dynamic information condition ($M_{Threat\ Rating} = 4.11$, $SD_{Threat\ Rating} = .69$) were notably higher than in the static information condition ($M_{Threat\ Rating} = 3.56$, $SD_{Threat\ Rating} = .53$, $t(38) = 2.81$, $p=.008$, $d=.89$). This suggests that all stimuli were perceived as more threatening when presented as point light walkers.

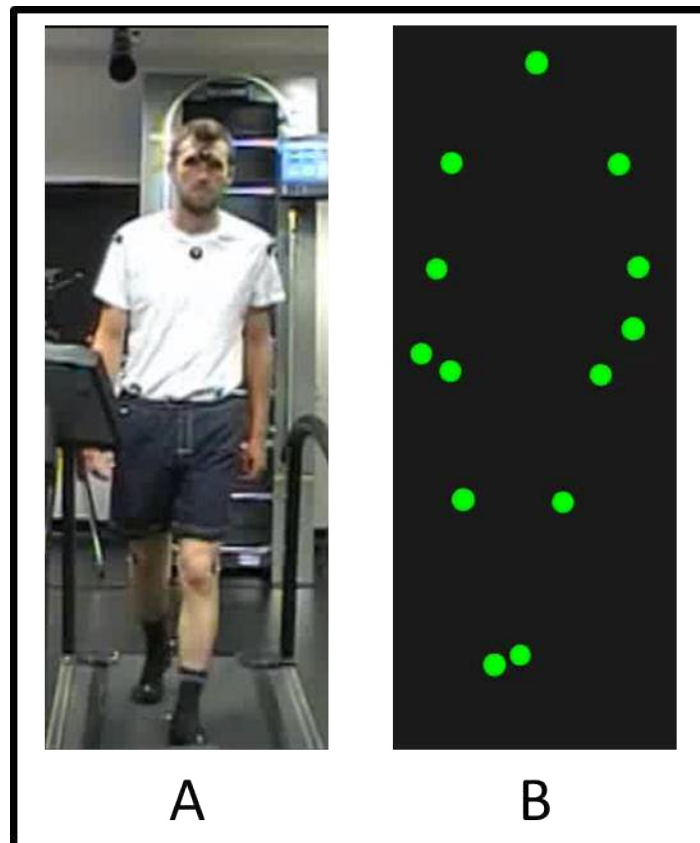


Figure 3.1. The appearance of targets in the two experiments. A) The presentation of targets in the first study and the second study's static information conditions. B) The point light walkers used as the dynamic information condition in the second study.

3.3.2.2 The relationship between threat ratings received and trait aggression.

Two correlations investigated the relationships between the average threat ratings received by the targets in each condition and the targets' self-reported aggression. There was no relationship between average threat ratings received by the targets and the targets' trait aggression in the static information alone condition ($r(23) = .38$, 95% CI $[-.07, .75]$, $p = .070$). In the dynamic information alone condition there was evidence of participants accurately detecting more aggressive targets in their threat ratings ($r(23) = .52$, 95% CI $[.10, .79]$, $p = .010$). Whilst there was no significant difference in the size of these correlations (using a Fisher's z test, $Z = .56$, $p = .580$), the dynamic information demonstrated a stronger relationship between threat judgments and targets' traits.

There was a similar pattern of results when the male targets and female targets were analysed separately. For male targets there was a stronger correlation between the average threat rating received by the targets and the target's trait aggression in the dynamic condition ($r(11) = .65$, 95% CI $[.14, .95]$, $p = .030$) than in the static condition ($r(11) = .52$, 95% CI $[-.17, .85]$, $p = .100$), but the size of correlation did not significantly differ ($Z = .40$, $p = .689$). For female targets there was a larger difference between accuracy for judging traits in the static condition ($r(12) = .31$, 95% CI $[-.10, .89]$, $p = .323$) and the dynamic condition ($r(12) = .61$, 95% CI $[-.55, .90]$, $p = .037$) but this, still, was not significantly different ($Z = .82$, $p = .412$). Whilst there were no differences between the size of the correlations in Fisher's Z tests, it is notable that only the accuracy correlations in the dynamic conditions are significant.

3.3.3 Study Two Summary Discussion. In the second study we demonstrated that threat judgments accurately predict target trait aggression when gait information is available to perceivers. When there was only static information available, threat judgments were not able to accurately reflect aggression. Whilst direct tests comparing accuracy in the dynamic and static conditions showed no significant differences, the dynamic condition accuracy values were higher than the static when the analysing the targets as a whole, just analysing the male targets and just analysing the female targets. These results suggest that the accuracy of threat judgments is best when perceivers have access to movement information.

3.4 General Discussion

The current paper investigated whether or not judgments of threat, perhaps informed by a ‘masculine is dangerous’ heuristic, could be accurate. Across two studies we found evidence that threat judgments made from full body, walking stimuli reflected the trait aggression of our targets. In the first study we also demonstrated the influence of the masculinity of body shape on threat judgments. It is somewhat surprising that the accuracy of participants’ threat ratings was so strong, as often there were very small numerical differences in the psychometric measure of aggression between targets.

The extant research that demonstrates the presence of the ‘masculine is dangerous’ heuristic typically uses faces as stimuli (Carré, et al., 2009; Hehman, et al., 2015; Hehman, et al., 2013) and our findings extend that research, showing that full body stimuli are susceptible to the same influences. More than this heuristic, the results of both studies showed that threat judgments received by the targets were accurate reflections of the targets’ trait aggression. This finding is interesting as it shows that the participants were accessing information relevant to the traits of the target and the targets appeared as threatening when they had self-rated themselves to be aggressive

In our second study we further investigated how accuracy in threat judgments could be possible by manipulating the availability of information from the targets. The Realistic Accuracy Model (Funder, 1990) states that accurate judgments of another's traits are only possible when the information relevant to a trait being judged are available for a judge to detect and utilize. When we limited the content of the study one videos so that only dynamic or only static information was available, we found that movement was more important for accurate threat judgments of potential aggressors than image. This was notable in the inaccurate threat judgments made by participants when movement was not available (in our static information condition).

The results presented here could be explained because movement is variable. Motion could be more reflective of personality than fixed body shape because the environment and experiences of an individual may shape their disposition, whereas endocrine effects on body shape cease to have large effects on body shape after puberty. The ability to manipulate one's motion could also allow a person to appear more or less of a threat in a given situation. By manipulating their walk targets may be able to mask their true traits. In the current study our targets were told to walk naturally. Even with a 'normal' walk, with no intention to appear more or less threatening, there was communication of trait aggression. In fact, research using real CCTV footage of impending crimes highlights the importance of the gait in detecting impending crimes (Troscianko et al., 2004).

Our measure of aggression in this experiment, the physical aggression scores of the Buss-Perry Aggression Questionnaire may have been limiting. Whilst the test has been shown to have external validity (Archer & Webb, 2006; Diamond, 2006; O'Connor, Archer & Wu, 2001) and the physical aggression measure is most relevant to our studies, future research could explore other measures of aggression. Although, it should be noted

that there is literature to suggest caution with using ‘experimental’ measures of aggression (Tedeschi & Quigley, 1996; Ritter & Eslea, 2005). In particular future works could explore how (or if) more social forms of aggression (such as verbal aggression or hostility) is communicated through movement or physical appearance.

3.4.1 Conclusion. Presented here, is evidence of threat judgments accurately reflecting the trait aggression of a walking target. The results also show that participants’ threat ratings were influenced by the assumption that the more masculine the body of the target, the more of a threat the target posed. These studies build a strong case for the importance of using dynamic, full body presentations of target stimuli in interpersonal judgments research. Further, these experiments provide theoretical support for work in detecting real street crime (such as; Trocianko et al., 2004).

References

- Archer, J. (1991). The influence of testosterone on human aggression. *British Journal of Psychology*, 82(1), 1–28. doi:10.1111/j.2044-8295.1991.tb02379.x
- Archer, J., & Webb, I. A. (2006). The relation between scores on the Buss–Perry Aggression Questionnaire and aggressive acts, impulsiveness, competitiveness, dominance, and sexual jealousy. *Aggressive Behavior*, 32(5), 464–473. doi:10.1002/ab.20146
- Bryant, F. B., & Smith, B. D. (2001). Refining the Architecture of Aggression: A Measurement Model for the Buss–Perry Aggression Questionnaire. *Journal of Research in Personality*, 35(2), 138–167. doi:10.1006/jrpe.2000.2302
- Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology*, 63(3), 452–459. doi:10.1037/0022-3514.63.3.452
- Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial Structure Is a Reliable Cue of Aggressive Behavior. *Psychological Science*, 20(10), 1194–1198. doi:10.1111/j.1467-9280.2009.02423.x
- Cheek, J. M. (1982). Aggregation, moderator variables, and the validity of personality tests: A peer-rating study. *Journal of Personality and Social Psychology*, 43(6), 1254–1269. doi:10.1037/0022-3514.43.6.1254
- Daly, M., & Wilson, M. (1990). Killing the competition. *Human Nature*, 1(1), 81–107. doi:10.1007/bf02692147
- Diamond, P. M. (2006). The Short-Form Buss-Perry Aggression Questionnaire (BPAQ-SF): A Validation Study With Federal Offenders. *Assessment*, 13(3), 227–240. doi:10.1177/1073191106287666
- Eagly, A. H., & Steffen, V. J. (1986). Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin*, 100(3), 309–330. doi:10.1037/0033-2909.100.3.309

- Funder, D. C. (1980). On seeing ourselves as others see us: Self-other agreement and discrepancy in personality ratings. *Journal of Personality*, 48(4), 473–493.
doi:10.1111/j.1467-6494.1980.tb02380.x
- Funder, D. C. (Ed.). (1999). *Personality judgment: A realistic approach to person perception*. Academic Press.
- Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, 21(3), 177-182.
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, 62, 451-482.
- Good, J.M.M. (2007). The affordances for social psychology of the ecological approach to social knowing. *Theory and Psychology*, 17, 265-295.
- Gray, J. P., & Wolfe, L. D. (1980). Height and sexual dimorphism of stature among human societies. *American Journal of Physical Anthropology*, 53(3), 441-456.
- Greitemeyer, T. (2015). Everyday sadism predicts violent video game preferences. *Personality and Individual Differences*, 75, 19-23.
- Gunns, R. E., Johnston, L., & Hudson, S. M. (2002). Psychopathic Traits and Perceptions of Victim Vulnerability. *Journal of Nonverbal Behavior*, 26(3), 129–158.
doi:10.1023/a:1020744915533
- Hehman, E., Flake, J. K., & Freeman, J. B. (2015). Static and Dynamic Facial Cues Differentially Affect the Consistency of Social Evaluations. *Personality and Social Psychology Bulletin*, 0146167215591495.
- Hehman, E., Leitner, J. B., & Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology*, 49(4), 747–754.
doi:10.1016/j.jesp.2013.02.015

- Hughes, S. M., & Gallup, G. G. (2003). Sex differences in morphological predictors of sexual behavior. *Evolution and Human Behavior*, 24(3), 173–178.
doi:10.1016/s1090-5138(02)00149-6
- Lake, S. L., Stanford, M. S., & Patton, J. H. (2014). Emotional processing and frontal asymmetry in impulsive aggressive individuals. *Personality and Individual Differences*, 70, 131-135.
- McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology*, 43(2), 293–303.
doi:10.1037/0022-3514.43.2.293
- McNaughton, N., & Corr, P. J. (2004). A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance. *Neuroscience & Biobehavioral Reviews*, 28(3), 285-305
- O'Connor, D. B., Archer, J., & Wu, F. W. C. (2001). Measuring aggression: Self-reports, partner reports, and responses to provoking scenarios. *Aggressive Behaviour*, 27(2), 79–101. doi:10.1002/ab.2
- Ritter, D., & Eslea, M. (2005). Hot Sauce, toy guns, and graffiti: A critical account of current laboratory aggression paradigms. *Aggressive Behavior*, 31(5), 407–419.
doi:10.1002/ab.20066
- Tedeschi, J. T., & Quigley, B. M. (1996). Limitations of laboratory paradigms for studying aggression. *Aggression and Violent Behavior*, 1(2), 163–177. doi:10.1016/1359-1789(95)00014-3
- Thoresen, J. C., Vuong, Q. C., & Atkinson, A. P. (2012). First impressions: Gait cues drive reliable trait judgements. *Cognition*, 124(3), 261–271.
doi:10.1016/j.cognition.2012.05.018

- Troscianko, T., Holmes, A., Stillman, J., Mirmehdi, M., Wright, D., & Wilson, A. (2004). What happens next? The predictability of natural behaviour viewed through CCTV cameras. *Perception*, 33(1), 87–101. doi:10.1068/p3402
- Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, 95(5), 1202–1216. doi:10.1037/a0013314
- Waldron, J. C., Scarpa, A., Lorenzi, J., & White, S. W. (2015). Depression mediates the relationship between social performance impairment and hostility. *Personality and Individual Differences*, 85, 165-171.
- Wilson, M., & Daly, M. (1985). Competitiveness, risk taking, and violence: the young male syndrome. *Ethology and Sociobiology*, 6(1), 59–73. doi:10.1016/0162-3095(85)90041-x
- Wilson, M., & Daly, M. (1993). Lethal confrontational violence among young men. In N. Bell & R. Bell (Eds.), *Adolescent risk taking* (pp. 84-106). Thousand Oaks, CA: US: Sage Publications.
- Zajenkowski, M., & Zajenkowska, A. (2015). Intelligence and aggression: The role of cognitive control and test related stress. *Personality and Individual Differences*, 81, 23-28.

4. Don't Look at the Face; Why the Body Matters for Detecting Trait Aggression

Manuscript in preparation for publication

Satchell, L., Nee, C., Morris, P., & Akehurst, L. (in preparation) *Don't look at the face;
why the body matters for detecting trait aggression*

Abstract

The objective of this study is to investigate how individuals' accuracy in detecting aggression in others is influenced by how that individual observes a target person. It is possible that more accurate judges of threat spend more time observing the aggression-relevant body movements of an approaching person. 27 judges observed 12 target individuals walking for 10 seconds and were tasked with judging how 'threatening' the target individual was. Participant accuracy at detecting trait aggression was calculated as the relationship between participants' threat ratings and the self-reported trait aggression of the respective targets. The percentage of time participants spent observing the targets' heads, bodies and legs was recorded using an eye tracker. We found that in the first 0 to 1 seconds of the video there is a strong tendency for all participants to observe the heads of the targets. However, in the following seconds it was the less accurate judges of aggression who continued to observe the face of the targets, whilst the more accurate judges preferred observing the body (and therefore walk) of the targets. We conclude that those who were more accurate judges of trait aggression were those who prioritised gait style over the face of the target.

4.1 Introduction

How might we accurately detect trait aggression in approaching people? Most research investigating perceptions of others' malevolent attributes uses static images of the face (such as; Carré, McCormick & Mondloch, 2009; Carré, Morrissey, Mondloch & McCormick, 2010; Geniole, Denson, Dixon, Carré & McCormick, 2015; Hehman, Leitner & Gaertner, 2013; Short et al, 2012). Static images, however, limit the amount of available (to use Funder's [1999] term) information for a perceiver to use in their judgment. Research has shown that the consistency and quality of judgments made of an individual can change when a rater has access to dynamic features of the face (Hehman, Flake & Freeman, 2015; Rubenstein, 2005). Presenting only the face of a target is also problematic. Observing faces involves fundamentally different neurological processes to observing the whole body (Kret, Pichon, Grèzes & de Gelder, 2011). Although some have claimed that faces alone communicate aggressive tendencies (Carré & McCormick, 2008) recent research (showing that facial shape is a less powerful predictor of trait aggression than bodies) has questioned this assumption of the communicative power of the face (Deaner, Goetz, Shattuck & Schnotala, 2012). Experiments using real CCTV footage of events leading to crimes have found that individuals are able to predict an impending crime from watching how an individual behaves, specifically from "distinctive gaits and hand gestures" (Troscianko et al., 2004, p.93). If distinctive gait is important for an observer to predict a crime, so could gait be used to detect aggression propensity from an approaching individual.

The body in motion communicates important information about emotions (Montepare, Goldstein & Clausen, 1987; Roether, Omlor, Christensen & Giese, 2009; Venutre, 2010), identity (Cutting & Kozlowski, 1977; Mather & Murdoch, 1994), vulnerability (Grayson & Stein, 1981; Gunns, Johnston & Hudson, 2002) and trait

aggression (Chapters 2, 3, 5 & 6). Most importantly, the whole body is the most typical way that individuals encounter others in the world (Azarian, Esser & Peterson, 2015b). In the case of observing another person, it is rare to only be exposed to the face of that individual and to have no access to that person's body, gait and general appearance. Judging the aggressiveness of an approaching person, therefore, could be influenced by different, previously unstudied, aspects of that approaching person. In our current work we investigate how judges of threat observe an approaching person when tasked with judging how threatening that person is.

Relatively accurate judgments about the threat posed by a target can be made after simply watching someone walking, without the target person engaging in overtly aggression-typical behaviours (Chapter 2). Whilst there is evidence that trait recognition, in general, is possible (Albright et al., 1988; Funder, 1980; 1999; 2012; McCrae, 1982), there is little research investigating how an individual's approach to observing a target relates to the recognition of that target's traits. It has been noted that eye tracking technology might offer a useful insight into unpacking strategies of observation for trait recognition (Hirschmüller, Egloff, Nestler, & Back, 2013).

Eye tracking in the context of threat detection has largely focused on photographs of faces. This research typically looks at the recognition of and fixation on 'threatening' emotional expressions, such as angry faces (e.g., Eastwood, Smilek & Merikle, 2001; Öhman, Lundqvist & Esteves, 2001) and averted eye gaze (e.g., Fox, Mathews, Calder & Yiend, 2007; Terburg, Hooiveld, Aarts, Kenemans, & van Honk, 2011). Some eye tracking research has investigated how individuals visually attend to the threatening postures of the whole body. However, these studies also use stimuli that are unrealistic for real-world threat detection, such as mannequins (Gilbert, Martin & Coulson, 2011) and exaggerated demonstrations of emotion (Azarian, Esser & Peterson 2015a; Azarian et al., 2015b).

Whilst this research contributes to the understanding of how threat is attended to with immediate risks (once someone moves to fight), these studies do not show us how patterns of visual search might facilitate the evaluation of the potential (not imminent) threats.

In this study, therefore, we investigate how strategies of visual search relate to accuracy in detecting threat from an approaching and potentially aggressive target. Given that trait aggression may be communicated through gait (Chapter 2), we would expect those judges who are better able to detect trait aggression to attend more often to the gait-relevant features of the target. However, there is evidence of a general preference for individuals to choose to observe faces in the world (Cerf, Harel, Einhäuser & Koch, 2008), but it is unclear whether watching the face aids or detracts from accurate threat detection. In this study we present videos of whole body movement and compare the visual tracking strategies of observers in relation to the accuracy of subsequent judgements of threat.

4.2 Method

4.2.1 Participants. 30 participants judged the threat of the targets. Those with poor eye tracking calibration (see Section 4.2.3, $N=2$) and those who did not engage with the judgment part of the experiment ($N=1$) were excluded from final analysis, leaving a sample of 27 participants ($M_{Age}=26.78$, $SD_{Age}=10.02$, Female=20). Participants were paid £5 in shopping vouchers for their time and were recruited from a participant database consisting of members of the public and university students and staff. Participants are henceforth referred to as ‘judges’ to avoid confusion with the targets.

4.2.2 Materials.

4.2.2.1 Target videos. We filmed 12 female targets ($M_{Age}=20.57$, $SD_{Age}=2.02$), who were blind to the aim of the experiment, as targets of the threat ratings. The aim of this exploratory study was to demonstrate any relationship between observing gait and detecting trait aggression. Therefore we used targets who were most likely to have cues to

aggression in their gait. Previous research shows that there is a particularly strong association between gait and trait aggression in females (Chapter 2), therefore we used only female targets. Thus, if judges used targets' gait in their judgments of threat then the effect would be most apparent with these targets, demonstrating the proposed effect. All targets self-reported being "White" or "Caucasian" ($M_{Age}=20.58$ years, $SD_{Age}=1.78$, $Min_{Age}=18$, $Max_{Age}=24$).

We filmed the targets walking on a treadmill wearing standardised clothes (grey vest top and black leggings). All videos were 10 seconds of uninterrupted gait towards the camera and targets displayed a neutral facial expression.

We measured target trait aggression using a self-reported measure of aggression (the Buss-Perry Aggression Questionnaire; Buss & Perry, 1992, analysed using revisions suggested by; Bryant & Smith, 2001, see Appendix F). The Buss-Perry Aggression Questionnaire is well established and cited tool in investigating trait aggression, with good evidence of external validity (Archer & Webb, 2006; Diamond, 2006; O'Connor, Archer & Wu, 2001). We use the Physical Aggression subscale for analysis, as it was most pertinent to the likelihood to commit acts of aggression. This subscale, when analysed using advised revisions (Bryant & Smith, 2001), gives a score between 3 and 21 and our twelve walkers were reasonably spread in the range of possible scores ($M_{Aggression}=7.50$, $SD_{Aggression}=4.89$, $Min_{Aggression}=3$, $Max_{Aggression}=19$).

4.2.2.2 Eye tracking. Judges' eye movements were tracked at 1000 Hz with the EyeLink 1000 (SR Research, Ltd, Osgoode Canada), using pupil and corneal reflection to detect gaze. Judges placed their heads on a chin-rest at 50cm from the screen where targets' videos were presented.

4.2.3 Procedure. Judges first had their eye movements calibrated on the eye tracker. Calibration involved measuring the difference between the expected and actual

fixation positions on a 9 point grid presented on the screen. Deviation greater than 0.50° was considered too imprecise and calibration was repeated until the eye movements were tracked with greater accuracy. If, after repeated calibration, deviation was still greater than 0.50° then that judge was excluded from analyses (see section 4.2.1).

After successful calibration, judges were told that they would watch a series of videos and be asked to rate the targets on 1-7 Likert scales; *threatening-non-threatening*, *masculine-feminine*, and *attractive-unattractive* (masculinity and attractiveness were include as typical distractor variables and are not analysed here). This video-then-rating sequence occurred 12 times so that all the judges saw all the targets. The presentation order of the targets was randomised for each judge.

4.2.4 Data Analyses. Each judge observed 3 fixation areas (head, trunk and legs) of 12 targets walking over 10 seconds and then judged the targets on how threatening the target appeared. We compared this threat rating to the targets' known aggression for a measure of accuracy in threat judgments (henceforth referred to simply as judge *accuracy*).

Each judge's accuracy, each target and each time period could have a different relationship with the amount of time spent observing each fixation area. We sought to use a comprehensive analysis of our data (whilst avoiding an overcomplicated repeated measures test on all these factors). As the interest of the paper is on the visual search of judges, we chose to average the observation across targets (to test the observation of the *average target*). Whilst this may lose some variation in observing different targets, it does allow for best analysis of how visual search changes over time and with judge accuracy. Thus, we investigated the percentages of time spent fixating on the different areas of the *average target*, and how judge accuracy affects judge observation of the average target.

4.2.4.1 Critical areas of fixation. The observation of the targets is split into the percentage of each second judges spend studying the 3 critical areas of the targets; the head

(neck and above), the trunk (between neck and hips) and the legs (hips and below). The width of these areas was standardised as the width of that target's shoulders. Any observation of the target not within a critical fixation area is considered non-target observation. In total the time spent observing the head, trunk, legs and non-target observation is equal to 100%.

4.2.4.2 Accuracy analyses. We calculated the accuracy of each judge. We did this by correlating each judge's threat ratings of the targets with the targets' trait physical aggression (following the advice of similar trait judgment papers, Hirschmüller, Egloff, Schmukle, Nestler & Back, 2015; Kolar, Funder & Colvin, 1996). This produced a Pearson r value for each judge. This acted as an index of accuracy, where a value of $r = 1$ indicates that the most aggressive targets were judged as the most threatening, a value of $r = -1$ indicates that the least aggressive targets were judged as the most threatening and a value of $r = 0$ indicates that there is no relationship between judge threat judgments and target trait aggression. We did not form categories of 'accurate' or 'inaccurate' judges as this weakens the richness of the degree of accuracy calculation we have.

4.3 Results

4.3.1 Observation style over time. Using a repeated measures ANOVA we tested for the difference in the amount of time spent observing the fixation areas (3; Head, trunk and legs) and for an interaction with fixation area over time (10 time periods). The amount of time spent observing the different fixation areas of the average target varied significantly, which is demonstrated in a main effect of fixation area ($F(1.16, 30.24) = 48.37, p < .001$)⁶. There was a preference across the sample for fixating on the head ($M_{\text{Observation}} = 50.42\%$, $SD_{\text{Observation}} = 5.44\%$), more than the trunk ($M_{\text{Observation}} = 31.03\%$, $SD_{\text{Observation}} = 2.21\%$, $p = .004$) and the legs ($M_{\text{Observation}} = 11.06\%$, $SD_{\text{Observation}} = 5.06\%$,

⁶ All ANOVA violate sphericity (all Mauchly's $W < .28$, all $\chi^2 > 31.80$, all $p < .001$) therefore Greenhouse-Geisser corrections are used.

$p < .001$) of the average target. It is also notable that the judges observed the trunk more than the legs of the average target ($p < .001$).

There was an interaction between the amount of time spent observing each critical fixation area of the average target across time⁷ ($F(4.90, 127.46) = 4.56, p = .001$), demonstrating that observation of the average target changed over time (see table 4.1).

It is notable that the variance in the amount of time spent fixating on the different areas (the standard deviation in observation time, see table 4.1) of the average target was large for all body areas, indicating large amounts of individual variation in observation style. As the length of the video increased, the less time was spent observing the head (as demonstrated in the correlation between number of seconds passed and the percentage of time spent observing the head of the average target is negative, $r(10) = -.70, p = .023$) and torso (although not significantly, $r(10) = -.55, p = .103$) and more time was spent fixating on the legs (although not significantly, $r(10) = .57, p = .087$). It is also important to note that over the course of the video judges spent an increased amount of time not observing the target ($r(10) = .92, p < .001$, although the sample never spent more than 10.80% of a second not observing the target).

Put simply, whilst the head was continually the point of most fixation for the sample, observation after the first second of the video was more distributed over the trunk and legs of the target. The legs were notably the least observed part of the target, but do receive notable attention between 3 and 5 seconds.

4.3.2 Observation style and accuracy. There was a reasonable spread of judges' ability to detect target trait aggression through threat ratings (judge *accuracy*) in the current study ($M_{Accuracy} = .16, SD_{Accuracy} = .22, Min_{Accuracy} = -.34, Max_{Accuracy} = .65$). To investigate how judge accuracy related to the visual search of targets, we correlated the

⁷ There was a main effect of time period in the ANOVA ($F(2.00, 51.92) = 3.33, p = .044$), but this is an artefact of brief non-target observation towards the end of the videos (at most 10.80% of a second [between 8 and 9 seconds] is spent not observing the target.) and is not considered relevant to report in main text.

judges' individual accuracy r value (treating judge accuracy as a trait of the judges) with the percentage of time the judges spent observing the head, trunk and legs of the average target. We do this for each time period (second) of the average target video. A positive correlation would suggest that a more accurate judge of aggression would spend more time fixating on that part of the target. A negative correlation would demonstrate that a less accurate judge of aggression would spend more time observing that part of the target.

We found that *less* accurate judges spent *more* time fixating on the head of the average target (see table 4.2). At all the time periods, all correlations between the percentage of time spent fixating on the head of the average target and judge accuracy were negative. The percentage of time spent observing the head over the whole video, also negatively correlated with judge accuracy. The time period that shows the greatest negative association between observation and accuracy is between 3 and 4 seconds, suggesting that this is the time period most influential on judges' threat ratings.

The percentage of time spent fixating on the trunk of the target follows the opposite pattern; *more* accurate judges spent *more* time fixating on the trunk of the target (see table 4.2). Nearly all correlations between trunk fixation and accuracy are positive, with the exception of the last second. The percentage of time spent observing the trunk over the whole video positively correlated with judge accuracy. The largest correlation between trunk fixation and judge accuracy is between 3 and 4 seconds.

The relationship between leg fixation and accuracy varied over the time periods (see table 4.2). The results show that the most accurate judges fixated on the legs between 1 and 2 seconds. But this preference declines quickly and by the 4th second the more accurate judges fixate elsewhere and the less accurate judges observe the legs. This changed at between 8 and 9 seconds where the more accurate judges are those who fixate on the legs more. The average amount of time spent observing the legs over the length of

Table 4.1. *The Average Percentage of Time Judges Spent Observing Each of the Body Areas of the Average Target (with Standard Deviation in Brackets).*

Time period (seconds)	Body Area of Target		
	Head	Trunk	Legs
0-1	63.16 (27.98)	30.08 (25.86)	0.73 (1.40)
1-2	57.12 (25.22)	34.14 (21.00)	3.89 (5.06)
2-3	50.21 (20.98)	34.71 (13.53)	10.43 (10.61)
3-4	46.52 (21.53)	31.38 (12.60)	16.13 (12.96)
4-5	47.80 (19.15)	28.72 (12.80)	16.91 (11.41)
5-6	48.69 (19.29)	30.72 (13.05)	13.56 (10.01)
6-7	46.96 (18.83)	32.53 (15.07)	12.32 (8.03)
7-8	46.62 (19.21)	30.84 (18.84)	12.24 (6.00)
8-9	49.14 (18.68)	28.40 (13.91)	11.67 (7.18)
9-10	47.93 (17.61)	28.78 (11.99)	12.71 (9.40)
Average Time Period	50.42 (5.44)	31.03 (2.21)	11.06 (5.06)

Note: Not accounted for in the table is the % of non-target observation. In all cases non target observation is minimal and is less than 10.08% of a second

Table 4.2. *The Correlations Between Judge Overall Accuracy and Percentage of Time Spent Fixating on the Differing Areas of the Average Target Over the Length of the Video.*

Time period (seconds)	Body Area of Target		
	Head	Trunk	Legs
0-1	-.18	.18	-.11
1-2	-.30	.21	.36
2-3	-.33	.15	.29
3-4	-.42	.39	.20
4-5	-.27	.30	-.05
5-6	-.28	.32	-.11
6-7	-.23	.27	-.15
7-8	-.18	.13	-.26
8-9	-.26	.18	.34
9-10	-.15	-.09	.29
Average Time Period	-.32	.25	.12

Note: For ease of interpretation correlations equal t or larger than .20 are highlighted in **Bold**. Correlations less than -.20 are highlighted in *Italics*

the whole video did positively correlate with accuracy but not strongly. The influence of leg fixation on accuracy is interesting to report but should not be given too much attention due to the general lack of interest in observing the legs by the whole sample (average observation was always <17% of a second, see table 4.1). These results demonstrate that, whilst less accurate judges continually prefer to observe the head of the average target, the more accurate judges fixate more on the whole body (the legs and trunk).

4.4 Discussion

This study investigated how individuals' abilities to detect trait aggression in others may relate to their visual searching of an approaching person. Previous research into perceptions of others has a tendency to force judges into viewing only the static face of the target (such as; Carré et al., 2009; Geniole & McCormick, 2015; Geniol et al., 2015; Hehman, et al., 2013; Short et al, 2012), despite research demonstrating the importance of a face (Hehman, Flake & Freeman, 2015; Rubenstein, 2005) and body (e.g. Grayson & Stein, 1981; Gunns, Johnston & Hudson, 2002, Chapter 2) in motion. In the current work we allowed judges to observe the whole of a walking target's body and asked them to judge the threat that the individual posed.

We found that the area of most fixation throughout the video was the head and face of the average target. Previous research shows that faces are of interest in the world and are preferentially attended to in natural environments (Cerf et al., 2008). Here we found that around half of the length of the average video was spent observing the head. This result suggests that there are benefits to investigating perceptions of faces, as faces are observed preferentially to other parts of the body. However, this finding also highlights that *only* half of the average video was spent observing the face. It would therefore be beneficial for future research to also use full body presentations of targets, and investigate to what extent judges use the face of targets in making interpersonal judgments beyond 'threat'.

Whilst the head was the point of most fixation, the amount of time spent fixating on the head decreased quickly as the video progressed. This decrease was apparent even after only one second had passed. The effect of ‘prolonged’ observation of a target is under researched in itself, and in the context of the extant literature on observing human stimuli, observation of 10 seconds is *prolonged* observation. Research with static faces has shown that judgments of others can be reliably made from presentations lasting 1 second or less (Ballew & Todorov, 2007; Willis & Todorov, 2006, for an overview see; Todorov, Olivola, Dotsch & Mende-Siedlcki, 2014). Judges can make reliable judgments of aggression even from brief presentations of faces (39ms; Carré et al., 2010). But in everyday life, another person does not only appear for 39ms. It is possible that the additional information from the rest of the body (including gait) during the 10 seconds of approach could support or contradict the judgment made from the early microseconds of presentation.

Importantly, the style of observation varied depending on the accuracy of the judge at detecting trait aggression. More accurate judges chose to observe the body of the target more than less accurate judges. Previous research has demonstrated that gait can communicate trait aggression information (Chapter 2) and that threat judgments best reflect trait aggression when judges only see gait (Chapter 3). This study, however, is the first to demonstrate that more accurate judges choose to fixate on gait-relevant areas of a target more often than less accurate judges. In fact the overall negative relationship between observing the head of the average target and judge accuracy implies that the less accurate judges were those who preferred to fixate on the head of the average target. This demonstration of judge idiosyncrasy in observing human stimuli is rarely noted in the literature, and future research could investigate how individual’s understanding of a target can relate to their search of that target.

In the current work we only use female targets of threat judgments. Whilst it is typical for men to elicit the concept of threat more than women, our results are based on the ability to discriminate between more and less aggressive targets through threat judgments. We found that judges were able to identify the more aggressive women as more threatening. Although male targets may receive higher ratings of threat (higher average rating on a Likert scale), we would expect the ability to accurately judge more aggressive targets as more threatening to relate to the observation style of male targets as well. Future research could use male targets but we have no reason to expect anything other than the same discretionary effects found here, just with higher baseline threat judgments.

Research into *who* can make accurate judgments of potential aggressors could help with the selection of CCTV operators. As Troscianko et al. (2004) demonstrated, it is possible for individuals to detect an impending crime from observing the behaviours of individuals over CCTV. Our current research demonstrates that certain individuals are better at predicting the propensity to aggression of unknown people than others, and as such, may be better at predicting an impending crime from viewing CCTV footage. It is also important to note that the ability to detect a potential aggressor is beneficial skill to any member of the general public in avoiding becoming a victim of a crime. Important lessons may be learnt from future research in this area, including improving personal safety, training of security personnel and border security officials.

4.4.1 Conclusion. This study demonstrates that judges who are better able to detect trait aggression from an approaching person observe that person in a different way to someone who is less able to detect trait aggression. Those who are more accurate in their threat judgments are those who observe the whole target (the trunk and the torso) more than simply the head of the target.

References

- Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology*, 55(3), 387–395. doi:10.1037/0022-3514.55.3.387
- Archer, J., & Webb, I. A. (2006). The relation between scores on the Buss–Perry Aggression Questionnaire and aggressive acts, impulsiveness, competitiveness, dominance, and sexual jealousy. *Aggressive Behavior*, 32, 464–473. doi:10.1002/ab.20146
- Azarian, B., Esser, E. G., & Peterson, M. S. (2015a). Watch out! Directional threat-related postures cue attention and the eyes. *Cognition and Emotion*, 1–9. doi:10.1080/02699931.2015.1013089
- Azarian, B., Esser, E. G., & Peterson, M. S. (2015b). Evidence from the eyes: Threatening postures hold attention. *Psychonomic Bulletin Review*. doi:10.3758/s13423-015-0942-0
- Ballew, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences*, 104(46), 17948–17953. doi:10.1073/pnas.0705435104
- Bryant, F. B., & Smith, B. D. (2001). Refining the Architecture of Aggression: A Measurement Model for the Buss–Perry Aggression Questionnaire. *Journal of Research in Personality*, 35, 138–167. doi:10.1006/jrpe.2000.2302
- Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology*, 63, 452–459. doi:10.1037/0022-3514.63.3.452
- Carré, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behavior in the laboratory and in varsity and professional hockey players.

- Proceedings of the Royal Society of London, Series B: Biological Sciences*, 275, 2651–2656. doi:10.1098/rspb.2008.0873
- Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial Structure Is a Reliable Cue of Aggressive Behavior. *Psychological Science*, 20(10), 1194–1198. doi:10.1111/j.1467-9280.2009.02423.x
- Carré, J. M., Morrissey, M. D., Mondloch, C. J., & McCormick, C. M. (2010). Estimating aggression from emotionally neutral faces: Which facial cues are diagnostic? *Perception*, 39(3), 356–377. doi:10.1068/p6543
- Cerf, M., Harel, J., Einhäuser, W., & Koch, C. (2008). Predicting human gaze using low-level saliency combined with face detection. In *Advances in neural information processing systems* (pp. 241-248).
- Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society.*, 9(5), 353–356. doi:10.3758/bf03337021
- Deaner, R. O., Goetz, S. M. M., Shattuck, K., & Schnotala, T. (2012). Body weight, not facial width-to-height ratio, predicts aggression in pro hockey players. *Journal of Research in Personality*, 46(2), 235–238. doi:10.1016/j.jrp.2012.01.005
- Diamond, P. M. (2006). The Short-Form Buss-Perry Aggression Questionnaire (BPAQ-SF): A Validation Study With Federal Offenders. *Assessment*, 13, 227–240. doi:10.1177/1073191106287666
- Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception & Psychophysics*, 63(6), 1004–1013. doi:10.3758/bf03194519

- Fox, E., Mathews, A., Calder, A. J., & Yiend, J. (2007). Anxiety and sensitivity to gaze direction in emotionally expressive faces. *Emotion*, 7(3), 478–486.
doi:10.1037/1528-3542.7.3.478
- Funder, D. C. (1980). On seeing ourselves as others see us: Self-other agreement and discrepancy in personality ratings. *Journal of Personality*, 48(4), 473–493.
doi:10.1111/j.1467-6494.1980.tb02380.x
- Funder, D. C. (Ed.). (1999). *Personality judgment: A realistic approach to person perception*. Academic Press.
- Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, 21(3), 177-182. doi: 10.1177/0963721412445309
- Geniole, S. N., Denson, T. F., Dixon, B. J., Carré, J. M., & McCormick, C. M. (2015). Evidence from Meta-Analyses of the Facial Width-to-Height Ratio as an Evolved Cue of Threat. *PLoS ONE*, 10(7), e0132726. doi:10.1371/journal.pone.0132726
- Gilbert, T., Martin, R., & Coulson, M. (2011). Attentional biases using the body in the crowd task: Are angry body postures detected more rapidly? *Cognition & Emotion*, 25(4), 700–708. doi:10.1080/02699931.2010.495881
- Grayson, B., & Stein, M. I. (1981). Attracting assault: Victims' nonverbal cues. *Journal of Communication*, 31(1), 68-75. doi 10.1111/j.1460-2466.1981.tb01206.x
- Gunns, R. E., Johnston, L., & Hudson, S. M. (2002). Psychopathic Traits and Perceptions of Victim Vulnerability. *Journal of Nonverbal Behavior*, 26(3), 129–158.
doi:10.1023/a:1020744915533
- Helman, E., Flake, J. K., & Freeman, J. B. (2015). Static and Dynamic Facial Cues Differentially Affect the Consistency of Social Evaluations. *Personality and Social Psychology Bulletin*, 41, 1123-1134. doi:10.1177/0146167215591495

- Helman, E., Leitner, J. B., & Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology*, 49(4), 747-754. doi:10.1016/j.jesp.2013.02.015
- Hirschmüller, S., Egloff, B., Nestler, S., & Back, M. D. (2013). The dual lens model: A comprehensive framework for understanding self–other agreement of personality judgments at zero acquaintance. *Journal of Personality and Social Psychology*, 104(2), 335–353. doi:10.1037/a0030383
- Hirschmüller, S., Egloff, B., Schmukle, S. C., Nestler, S., & Back, M. D. (2015). Accurate Judgments of Neuroticism at Zero Acquaintance: A Question of Relevance. *Journal of Personality*, 83(2), 221–228. doi:10.1111/jopy.12097
- Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of personality*, 64, 311-338. doi:10.1111/j.1467-6494.1996.tb00513.x
- Kret, M. E., Pichon, S., Grèzes, J., & Gelder, B. de. (2011). Men Fear Other Men Most: Gender Specific Brain Activations in Perceiving Threat from Dynamic Faces and Bodies – An fMRI Study. *Frontiers in Psychology*, 2. doi:10.3389/fpsyg.2011.00003
- Mather, G., & Murdoch, L. (1994). Gender Discrimination in Biological Motion Displays Based on Dynamic Cues. *Proceedings of the Royal Society B: Biological Sciences*, 258(1353), 273–279. doi:10.1098/rspb.1994.0173
- McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology*, 43(2), 293–303. doi:10.1037/0022-3514.43.2.293
- Montepare, J. M., Goldstein, S. B., & Clausen, A. (1987). The identification of emotions from gait information. *Journal of Nonverbal Behavior*, 11(1), 33–42. doi:10.1007/bf00999605

- O'Connor, D. B., Archer, J., & Wu, F. W. C. (2001). Measuring aggression: Self-reports, partner reports, and responses to provoking scenarios. *Aggressive Behaviour*, 27, 79–101. doi:10.1002/ab.2
- Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*, 80(3), 381–396. doi:10.1037/0022-3514.80.3.381
- Roether, C. L., Omlor, L., Christensen, A., & Giese, M. A. (2009). Critical features for the perception of emotion from gait. *Journal of Vision*, 9(6), 15–15. doi:10.1167/9.6.15
- Rubenstein, A. J. (2005). Variation in Perceived Attractiveness: Differences Between Dynamic and Static Faces. *Psychological Science*, 16(10), 759–762. doi:10.1111/j.1467-9280.2005.01610.x
- Short, L. A., Mondloch, C. J., McCormick, C. M., Carré, J. M., Ma, R., Fu, G., & Lee, K. (2012). Detection of propensity for aggression based on facial structure irrespective of face race. *Evolution and Human Behavior*, 33(2), 121–129. doi:10.1016/j.evolhumbehav.2011.07.002
- Terburg, D., Hooiveld, N., Aarts, H., Kenemans, J. L., & van Honk, J. (2011). Eye Tracking Unconscious Face-to-Face Confrontations: Dominance Motives Prolong Gaze to Masked Angry Faces. *Psychological Science*, 22(3), 314–319. doi:10.1177/0956797611398492
- Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2014). Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annual Review of Psychology*, 66(1), 519–545. doi:10.1146/annurev-psych-113011-143831

Troscianko, T., Holmes, A., Stillman, J., Mirmehdi, M., Wright, D., & Wilson, A. (2004).

What happens next? The predictability of natural behaviour viewed through CCTV cameras. *Perception*, 33(1), 87–101. doi:10.1068/p3402

Venture, G. (2010). Human characterization and emotion characterization from gait.

Annual International Conference of the IEEE Engineering in Medicine and Biology.
doi:10.1109/iembs.2010.5626404

Willis, J., & Todorov, A. (2006). First Impressions: Making Up Your Mind After a 100-Ms

Exposure to a Face. *Psychological Science*, 17(7), 592–598. doi:10.1111/j.1467-9280.2006.01750.x

5. Learning to be Streetwise: The Acquisition of Accurate Judgments of Aggression

Based on the manuscript in press with *Psychiatry, Psychology & Law*

Satchell, L., Akehurst, L. & Morris, P. (in press) Learning to be streetwise: The acquisition of accurate judgments of aggression.

Psychiatry, Psychology and Law.

Abstract

The detection of a potential danger is an important factor in avoiding harm. This is even more important for vulnerable populations such as children. Based on previous evidence that adults are able to accurately identify potential aggressors, we explored whether children could recognise the potential for a dangerous encounter in a similar way. Participants were divided into three age groups; over 18 year olds, 16-17 year olds and 13-15 year olds. Participants made judgments of nine, point light presentations of people walking on a treadmill. Ratings of intimidation made by participants were used to assess their ability to detect the walkers' trait aggression. Only the intimidation ratings made by the participants over the age of 18 years correlated with the self-reported aggression of the walkers. There was a reasonable level of agreement between participants in judgments of level of intimidation of walkers; however, the level of agreement was higher within the older age groups. There were also positive correlations within all age groups between the masculinity of walkers and judgments of intimidation. The ability to accurately detect trait aggression in others increased with age as did the consistency in ratings between individuals within the same age group. We highlight the importance of experiential learning in the acquisition of aggression detection skills.

5.1 Introduction

Children are frequently told to avoid strangers in general, but there is evidence that young children learn the ‘stranger danger’ message without putting it into practice, even in a laboratory setting (Moran, Warden, Macleod, Mayes & Gillies, 1997). It is difficult for a child to develop the ability to avoid dangerous adults as most children are not routinely exposed to high risk street danger. Children are usually accompanied by a parent who will make judgments about a potential threat. In fact, becoming a parent heightens adults’ perceptions of the formidability of unknown people (Fessler, Holbrook, Pollack, & Hahn-Holbrook, 2014). This suggests that children have even less of a responsibility to detect potential aggressors as their parents will be overly cautious on their behalf. Over the past two decades, parents’ fear of strangers has led to less freedom for children (Foster, Villanueva, Wood, Christian, & Giles-Corti, 2014). For example, there has been an increase in the percentage of children accompanied on journeys to and from primary and secondary school (Shaw et al., 2013). With children becoming increasingly sheltered by their parents, it is of interest to investigate the ability of young people to detect potential aggressors.

Given the importance of efficiency in detecting potential dangers, it is no surprise that adults make judgments based on simple heuristics such as ‘masculine is dangerous’. This heuristic is well established in research that reports a relationship between facial masculinity and judgments of malevolent attributes (Carré, McCormick & Mondloch, 2009; Hehman, Flake & Freeman, 2015; Hehman, Leitner & Gaertner, 2013). What is rated as more masculine appears more dangerous, aggressive and intimidating.

Past research into judgments of danger has primarily focussed on static images of faces (e.g. Carré, et al., 2009; Hehman, et al., 2015; Hehman, et al, 2013). In reality, judgments are likely to start being formed at a distance meaning that facial cues are absent.

At this range, other information will be available such as body shape and the movements of a potential aggressor (Chapter 3). In fact, research has demonstrated that gait is an important factor in making accurate judgments of trait aggression (Chapter 2).

Funder's (1999) Realistic Accuracy Model (RAM) suggests that a target person affords a perceiver *relevant* and *available* cues to *detecting* personality traits. As noted above, this programme of research has shown that gait is a *relevant* and readily available cue to trait aggression. The *availability* of gait to a perceiver is notable as walking behaviours can be seen even in impoverished situations including at a distance or at night. This allows a judgment of a potential aggressor to be made earlier (e.g. in situations when time for avoidance behaviour is, or is at least perceived to be, limited, McNaughton & Corr, 2004).

However, as the RAM suggests, different perceivers may *utilize* (Funder, 1999) these cues in different ways. Costall (1995) notes that the understanding of affordances (in this case interpersonal affordances) may be culturally driven; "a child, for example, is not simply left to 'discover' the function of a cup or spoon; rather the learning situation involves careful structuring" (p.472). We suggest that Costall's comments on the social and cultural understanding of affordances from objects could also be used to understand age and cultural effects on the utilization of interpersonal affordances in the RAM. Just as a child learns the meaning of a spoon in a structured environment, so might the child learn the risk factors in the gait of an adult. We were interested in how the same gait information is utilized, by perceivers of different ages, to make intimidation judgments. These judgments may or may not reflect the walkers' trait aggression. Funder (2012) refers to interpersonal judgments that relate to the self-reported traits of another as 'accurate' judgments, and we adopt the same term here.

We chose to use point light walkers as stimuli as they have been shown to be effective at communicating trait aggression (see Chapter 2). This format of stimuli allows us to present gait cues and some body shape information, without facial characteristics or clothing of targets influencing judgments. Children as young as 5 years old have been shown to recognise point light stimuli as human (Pavlova, Krägeloh-Mann, Sokolov, & Birbaumer, 2001) and children as young as 35 months have been able to correctly identify point light motion as gait (Golinkoff et al, 2002) demonstrating that even the youngest participants in our sample (13 years) were able to engage with the task.

We explored the use of a ‘masculine is dangerous’ heuristic on participants’ judgments of intimidation. It was expected that all participants would use the ‘masculine is dangerous’ heuristic with more masculine bodies judged as more intimidating. However, the main aim of the current study was to investigate the effect that age had on the accuracy of intimidation ratings in detecting trait aggression. Whilst we expected adults would outperform younger participants, we had no firm expectations in terms of the younger participants’ accuracy.

5.2 Method

5.2.1 Participants. Eighty-five school children (46 females) were recruited for an ‘under 16 years’ age group (13 years to 15 years, $M_{Age} = 14.40$ years, $SD_{Age} = .73$ years). One hundred and three college students (82 females) were recruited for a ‘16-17 years’ age group ($M_{Age} = 16.42$ years, $SD_{Age} = .50$ years). Fifty-four undergraduate university students were recruited for an ‘over 18 years’ age group (44 females, $M_{Age} = 21.15$ years, $SD_{Age} = 3.21$ years).

5.2.2 Materials. This experiment used the same stimulus walkers as a previous study (Chapter 2). The walkers were presented in point light format. They were filmed with retroreflective markers placed on both heels, both knees, both greater trochanters (in

effect, hips), both shoulders, both elbows, both wrists and their foreheads. We filmed the individuals walking at their preferred speed for 10 seconds using Qualysis point light filming technology and ten ProReflex infrared cameras (100 Hz). This allowed us to present the walkers as 13 green dots on a black background.

A Sexual Dimorphism Index (SDI) score was calculated for each walker using their biometric information. The SDI score was a measure of how masculine-typical the walkers' bodies were relative to the other walkers in our study. The SDI scores were derived as a function of three sexually dimorphic features (waist-to-hip ratio, shoulder-to-hip ratio and height) and comprised a ranked score of 1-23. A higher SDI score implied more typically masculine morphology. All participants with an SDI score lower than 13 were female.

For the purpose of keeping the experiment efficient and retaining the engagement of the younger participants, a sub-set of nine walkers (from the database of 23 created for chapter 3) were chosen for use in this study. These nine walkers were chosen to represent different ends of the SDI spectrum due to the previously observed relationship between masculinity and perceptions of danger (Carré et al., 2009; Hehman, et al., 2015; Hehman, et al., 2013; Chapter 2). Three walkers were chosen as the most masculine (all male, $SDI \geq 21$), three as the most feminine (all female, $SDI \leq 3$) and three were the median-most values (two female, $SDI = 11$ and 12 and one male, $SDI = 13$).

Walkers completed the Buss-Perry Aggression Questionnaire (Buss & Perry, 1992), which we analysed using revisions suggested by Bryant and Smith (2001, see Appendix F). This questionnaire was used as it is a well-established measure of participants' tendencies to aggression. The questionnaire has been shown to be a valid measure of current aggression (Bryant & Smith, 2001), aggression in a hypothetical context (Archer & Webb, 2006; O'Connor, Archer & Wu, 2001) and historic aggression

(Diamond, 2006) and has been used with both student (García-León et al., 2002) and forensic (Diamond, Wang & Buffington-Vollum, 2005) populations. We only used the ‘trait physical aggression’ measure of the Buss-Perry Aggression Questionnaire as this is most relevant to interpersonal threat judgments. In accordance with Bryant and Smith’s (2001, see Appendix F) revisions, participants could score between 3 and 21 for the physical aggression measure and our sample of 9 undergraduate student walkers were surprisingly well spread within this range of potential scores ($M_{\text{Aggression}}=7.11$, $SD_{\text{Aggression}}=5.09$, $\text{Min}_{\text{Aggression}}=3$, $\text{Max}_{\text{Aggression}}=15$).

5.2.3 Procedure. The younger participants took part in small groups (sessions of 15-29 participants held in schools and colleges) and the older participants in one larger group (54 participants in a lecture setting.) The order of presentation of the nine walkers was randomised for each group. After the presentation of each walker, the participants were given as much time as they required to make ratings, on 9-point Likert scales, of how *intimidating-not intimidating*, *friendly-unfriendly* and *masculine-feminine* they thought each walker was.

5.2.4 Analyses. Using the procedures described in previous publications (Brand & Bradley, 2012; Kolar, Funder & Colvin, 1996; Monin & Oppenheimer, 2005) we analysed our data in two ways, using correlated averages and average correlations. Firstly, we analysed the correlation between average intimidation rating received by the walker and the body shape and trait aggression of each walker. This allowed for examination of the accuracy of the age groups in judging a potential aggressor as intimidating (see the results section ‘Ratings received by walkers’). Using this analysis is informative as it provides information about the collective accuracy of the sampled populations. However, detail is lost in terms of individual variation in accuracy. Furthermore, the statistical N for analysis is reduced to 9 (as the properties of 9 walkers are correlated with each other), thus not

being reflective of the sample size used in this research. Therefore, we also report the accuracy of our individual participants (as measured by Pearson's r values). We calculated participant accuracy by correlating the nine intimidation ratings made by each participant (relating to each of the nine walkers) with the self-reported aggression of the nine respective walkers. This individual accuracy correlation can be interpreted like any correlation, with a score of 1 demonstrating high accuracy (e.g. more aggressive targets are perceived to be more intimidating), a score of -1 demonstrating high inaccuracy (e.g. more aggressive targets are perceived to be less intimidating) and a score of 0 demonstrating no relationship between trait aggression and intimidation ratings (random performance).

It was therefore possible to examine the distribution of participant accuracy across the sample and the differences in accuracy abilities across age groups (see the results section 'Accuracy of participants'). Reporting the results for collective age groups and at an individual level provides a more complete understanding of both the properties of our walkers and the judgments of our participants.

5.3 Results

5.3.1 Ratings received by walkers. In this part of the results section we focus on the walkers; using the average intimidation rating received by each walker, each walker's trait aggression and each walker's body masculinity (SDI score). The SDI ranking of the walkers positively correlated with the average intimidation rating received by the walkers for all participant age groups (13-15 years, $r(9)=.78$, 95% CI [.43, .96], $p=.014$; 16-17 years, $r(9)=.77$, 95% CI [.34, 1.00], $p=.015$; over 18 years, $r(9)=.89$, 95% CI [.65, .99], $p=.001$) with very strong effects. This implies that in all cases there was a 'masculine is dangerous' heuristic being used by participants to judge the walkers' intimidation. This heuristic was reasonably accurate in itself for this set of walkers, with body masculinity

and trait aggression positively correlating with a medium effect ($r(9)=.64$, 95% CI [.04, .95], $p=.064$).

The average intimidation ratings the walkers received from the younger samples did not significantly correlate with the walkers' trait aggression (13-15 years, $r(9)=.36$, 95% CI [-.24, .77], $p=.341$; the 16-17 years, $r(9)=.40$, 95% CI [-.16, .91], $p=.281$).

However, the average ratings received by the walkers from the over 18 year olds were positively correlated with trait aggression with a large effect ($r(9)=.70$, 95% CI [.12, .97], $p=.036$). It should be noted that this judgment did not benefit from simply judging male targets as more intimidating as there was no effect of walker sex on walker aggression ($U(N=9)=5.00$, $p=.209$, see figure 5.1).

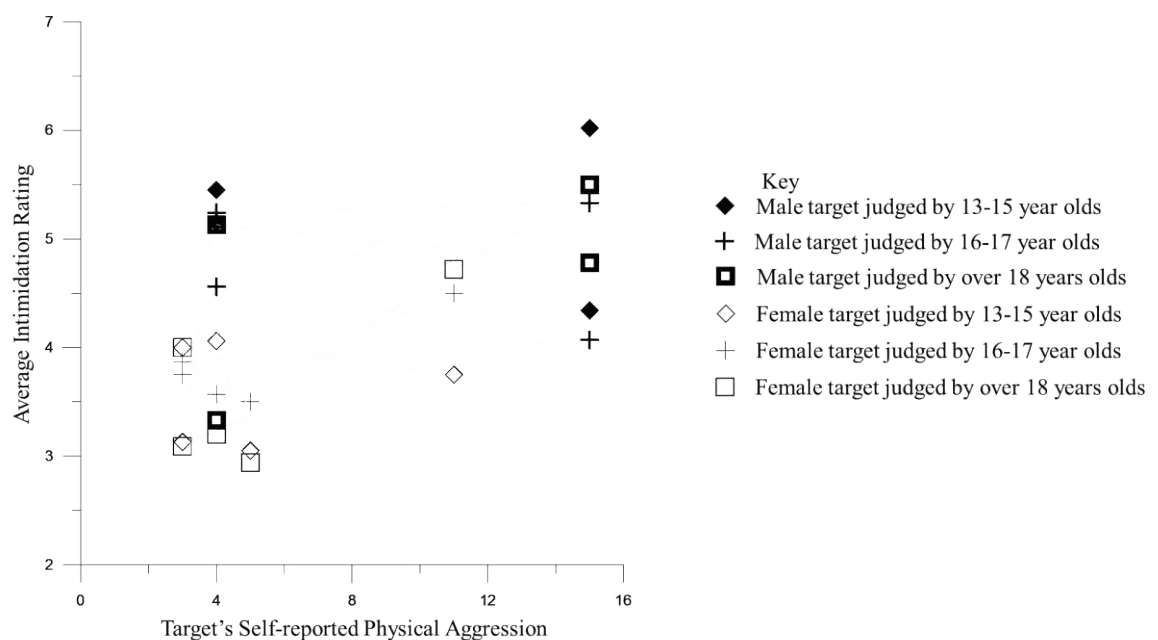


Figure 5.1 The relationship between average judgment of intimidation received by the walkers and the walker's trait aggression. Results are presented by age group and sex of walker.

It is interesting to note the variance in intimidation judgments made by participants in each of the age groups. Consistency in judgments received by the walkers reflects participants rating a construct they could easily identify. Variance in judgments

demonstrates more guess work or a lack of consensus about the matter being judged. By calculating the average standard deviation (σ) in intimidation ratings received by each walker we were able to test for differences in spread of intimidation judgments between the age groups. We found that the variation in intimidation ratings differed between age groups ($F(2,16)=8.81, p=.003, f^2=.70$) with the responses given by the under 16 year olds being more varied (having the highest σ , $M_\sigma=2.02, SD_\sigma=.10$) than those given by the 16-17 year olds ($M_\sigma=1.88, SD_\sigma=.12, p=.046, d=1.32$) and those given by the over 18 year olds (having the lowest σ , $M_\sigma=1.73, SD_\sigma=.22, p=.003, d=1.75$). The over 18 year olds and 16-17 year olds did not differ meaningfully ($p=.09, d=.89$). These results imply that a social consensus in judgments of intimidation develops with age.

5.3.2 Accuracy of participants. In this part of the results section we draw comparisons between the three age groups, using the accuracy correlations (r) of each participant as dependent variables (where 1 is perfect accuracy, -1 perfect inaccuracy and 0 random responding). On average, all age groups only achieved a small to medium level of accuracy ($M_r=.20, 95\% CI [.15, .23]$). See figure 5.2A for the distribution of the whole sample's performance. In fact, a binomial test showed that the majority of 13-15 year olds

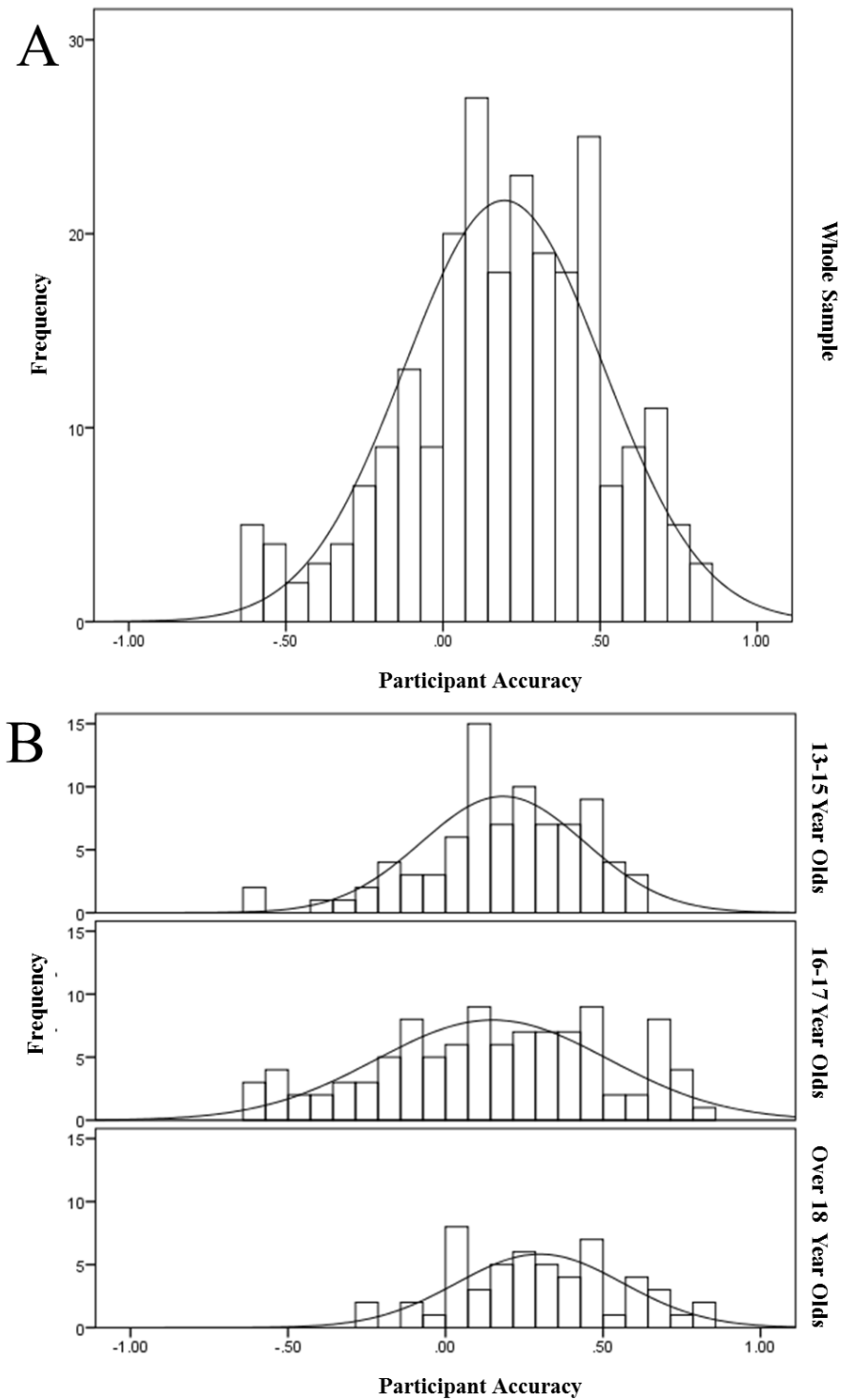


Figure 5.2. The distribution of participant accuracy in detecting trait aggression with intimidation ratings. Figure 5.1A demonstrates the overall sample accuracy and 5.1B separates out the participants by age category.

(81%, $p < .001$), 16-17 year olds (66%, $p = .001$) and over 18 year olds (89%, $p < .001$) had an accuracy correlation greater than zero, demonstrating overall accuracy. Few participants were notably inaccurate, with only 27 participants (11.16% of the whole sample) having an accuracy value less than $-.20$ (19 of whom were in the 16-17 year old condition). There was no overall correlation between age of participant and accuracy ($r(241) = .11$, 95% $CI [- .01, .23]$, $p = .11$). However, there were differences in accuracy when accuracy correlations were compared across age groups ($F(2, 238) = 4.18$, $p = .016$, $f = .19$). The over 18 year olds ($M_r = .30$, 95% $CI_r [.23, .37]$) were, on average, the most accurate in their intimidation ratings (more accurate than the under 16 year olds ($M_r = .15$, 95% $CI_r [.08, .22]$, $p = .01$, $d = .45$) and the 16-17 year olds ($M_r = .18$, 95% $CI_r [.13, .24]$, $p = .004$, $d = .45$). The 16-17 year olds and 13-15 olds performed similarly ($p = .47$, $d = .10$), see figure 5.2B for the distribution of performance by age categories.

5.4 Discussion

It is important for the recognition of potential aggressors to happen not only quickly but also accurately. It is possible that younger people do not have the life experience to detect potential dangers. In the current study we explored how accuracy in detecting potential aggressors develops with age. Participants were shown point light displays of targets. No facial characteristics or clothing were present in these videos, however gait and body shape information was available. All our participants, regardless of age, assumed that what is ‘masculine is dangerous’ and rated walkers with more masculine-typical bodies as more intimidating. When evaluating the accuracy of our participants we found that the older participants outperformed the younger participants in terms of their ability to detect trait aggression. In fact, all participants under the age of 17 performed similarly and were less accurate than those even a couple of years older (the over 18s group). Out of interest we analysed the spread of intimidation ratings. As the age

of participants increased, there was a decrease in the amount of variation in the judgments they made of the walkers.

These findings suggest that engaging in a more adult lifestyle exposes individuals to more situations which allow for aggressor detection skills to develop. Furthermore, with age, comes a social consensus regarding perceptions of intimidation. Importantly, the utilisation of this consensus seems to benefit the identification of potential aggressors. These findings contribute to our argument that accurate judgments of aggressors may be acquired through the socialisation of interpersonal affordances. Individuals' *utilization* (Funder, 1999) of gait affordances to make an accurate intimidation judgment may develop through engagement with a social world (Costall, 1995).

Our study used an under researched population; 13 to 17 year olds. There is a lack of interpersonal perception research on this *in-between* population who are not typically recruited in studies of children and are younger than the typically studied adult population (e.g. Cheek, 1982; Funder, 2012; McCrae, 1982; Vazire & Mehl, 2008). Research on interpersonal perceptions of teenagers almost entirely focuses on bullying (Boulton & Smith, 1994; Mynard & Joseph, 1997) and does not investigate *accuracy* of trait recognition. It is clear from the individual variation in accuracy for our under 18 year olds that although, as a group, they were worse at identifying potential aggressors than their adult counterparts, some of the younger participants could detect trait aggression in others. In fact some teenagers outperformed some adults. Future research should further investigate the factors that influence the acquisition of interpersonal accuracy.

Past research has fundamentally ignored the child as an active participant in judging danger. Studies have focused only on parental perceptions of formidability (Fessler et al., 2014) and their choice to supervise their children closely (Foster et al., 2014). It is important to study how children acquire the skills to detect legitimate dangers.

When out with their parents, children may well contribute to the decision to avoid or approach an unknown (or even known) person. In the current study, instead of considering parents' perceptions of their children's ability to 'manage' danger, we asked children themselves to make judgments. Further research should extend this methodology and should consider children as active participants in detecting dangers. Observations of joint decision making between children and their parents might reveal how parents teach their children to recognise potential aggressors. This would add further weight to our finding that children's acquisition of this ability develops over time.

The current study presented participants with people in motion using point light displays. It seems that gait is informative of many features of an individual, including trait aggression (Chapter 2, Chapter 3). Whilst previous research has focused on static facial stimuli (Carré et al., 2009; Hehman, et al., 2015; Hehman, et al., 2013), the current work demonstrates that gait alone is communicative enough of trait aggression. Gait is frequently *available* (Funder, 1999) from an approaching person and should be considered part of the overall impression when investigating aggression detection in real world contexts.

5.4.1 Conclusion. The current work has strong theoretical and applied implications. We demonstrated that the ability to use intimidation ratings to discriminate between more and less aggressive individuals is acquired with age. It seems that children learn to recognise the traits of others as they gain life experience. As children growing up in the 21st century are more sheltered than ever before by their parents it is interesting that not until the age of 17, when venturing further afield, do children become better able to accurately detect aggressors.

References

- Archer, J., & Webb, I. A. (2006). The relation between scores on the Buss–Perry Aggression Questionnaire and aggressive acts, impulsiveness, competitiveness, dominance, and sexual jealousy. *Aggressive Behavior*, 32, 464–473.
doi:10.1002/ab.20146
- Boulton, M. J., & Smith, P. K. (1994). Bully/victim problems in middle school children: Stability, self-perceived competence, peer perceptions and peer acceptance. *British journal of developmental psychology*, 12, 315-330. doi:10.1111/j.2044-835x.1994.tb00637.x
- Brand, A., & Bradley, M. T. (2012). More voodoo correlations: when average-based measures inflate correlations. *The Journal of general psychology*, 139(4), 260-272.
doi:10.1080/00221309.2012.703711
- Bryant, F. B., & Smith, B. D. (2001). Refining the Architecture of Aggression: A Measurement Model for the Buss–Perry Aggression Questionnaire. *Journal of Research in Personality*, 35, 138–167. doi:10.1006/jrpe.2000.2302
- Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology*, 63, 452–459. doi:10.1037/0022-3514.63.3.452
- Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial Structure Is a Reliable Cue of Aggressive Behavior. *Psychological Science*, 20, 1194–1198.
doi:10.1111/j.1467-9280.2009.02423.x
- Cheek, J. M. (1982). Aggregation, moderator variables, and the validity of personality tests: A peer-rating study. *Journal of Personality and Social Psychology*, 43, 1254–1269. doi:10.1037/0022-3514.43.6.1254
- Costall, A. (1995). Socializing affordances. *Theory & Psychology*, 5(4), 467-481.
doi:10.1177/0959354395054001

- Diamond, P. M. (2006). The Short-Form Buss-Perry Aggression Questionnaire (BPAQ-SF): A Validation Study With Federal Offenders. *Assessment*, 13, 227–240.
doi:10.1177/1073191106287666
- Diamond, P. M., Wang, E. W., & Buffington-Vollum, J. (2005). Factor structure of the Buss-Perry Aggression Questionnaire (BPAQ) with mentally ill male prisoners. *Criminal Justice and Behavior*, 32, 546–564.
doi:10.1177/0093854805278416
- Fessler, D. M. T., Holbrook, C., Pollack, J. S., & Hahn-Holbrook, J. (2014). Stranger danger: Parenthood increases the envisioned bodily formidability of menacing men. *Evolution and Human Behavior*, 35, 109–117.
doi:10.1016/j.evolhumbehav.2013.11.004
- Foster, S., Villanueva, K., Wood, L., Christian, H., & Giles-Corti, B. (2014). The impact of parents' fear of strangers and perceptions of informal social control on children's independent mobility. *Health & place*, 26, 60–68.
doi:10.1016/j.healthplace.2013.11.006
- Funder, D. C. (Ed.). (1999). *Personality judgment: A realistic approach to person perception*. Academic Press.
- Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, 21, 177–182. doi:10.1177/0963721412445309
- García-León, A., Reyes, G. A., Vila, J., Pérez, N., Robles, H., & Ramos, M. M. (2002). The Aggression Questionnaire: A validation study in student samples. *The Spanish Journal of Psychology*, 5, 45–53. doi:10.1017/s1138741600005825
- Golinkoff, R. M., Chung, H. L., Hirsh-Pasek, K., Liu, J., Bertenthal, B. I., Brand, R., ... Hennon, E. (2002). Young children can extend motion verbs to point-light displays. *Developmental Psychology*, 38, 604–614. doi:10.1037/0012-1649.38.4.604

- Hehman, E., Flake, J. K., & Freeman, J. B. (2015). Static and Dynamic Facial Cues Differentially Affect the Consistency of Social Evaluations. *Personality and Social Psychology Bulletin*, 41, 1123-1134. doi:10.1177/0146167215591495
- Hehman, E., Leitner, J. B., & Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology*, 49(4), 747-754. doi:10.1016/j.jesp.2013.02.015
- Hirschmüller, S., Egloff, B., Schmukle, S. C., Nestler, S., & Back, M. D. (2015). Accurate judgments of neuroticism at zero acquaintance: A question of relevance. *Journal of personality*, 83(2), 221-228. doi:10.1111/jopy.12097
- Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of personality*, 64, 311-338. doi:10.1111/j.1467-6494.1996.tb00513.x
- McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology*, 43, 293-303. doi:10.1037/0022-3514.43.2.293
- McNaughton, N., & Corr, P. J. (2004). A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance. *Neuroscience & Biobehavioral Reviews*, 28, 285-305. doi:10.1016/j.neubiorev.2004.03.005
- Monin, B., & Oppenheimer, D. M. (2005). Correlated averages vs. averaged correlations: Demonstrating the warm glow heuristic beyond aggregation. *Social Cognition*, 23(3), 257-278. doi:10.1521/soco.2005.23.3.257
- Moran, E., Warden, D., Macleod, L., Mayes, G., & Gillies, J. (1997). Stranger Danger: What do Children Know?. *Child Abuse Review*, 6, 11-23. doi: 10.1002/(SICI)1099-0852(199703)6:1<11::AID-CAR286>3.0.CO;2-G

- Mynard, H., & Joseph, S. (1997). Bully/victim problems and their association with Eysenck's personality dimensions in 8 to 13 year olds. *British Journal of Educational Psychology*, 67(1), 51-54. doi: 10.1111/j.2044-8279.1997.tb01226.x
- O'Connor, D. B., Archer, J., & Wu, F. W. C. (2001). Measuring aggression: Self-reports, partner reports, and responses to provoking scenarios. *Aggressive Behaviour*, 27, 79–101. doi:10.1002/ab.2
- Pavlova, M., Krägeloh-Mann, I., Sokolov, A., & Birbaumer, N. (2001). Recognition of point-light biological motion displays by young children. *Perception*, 30, 925–933. doi:10.1068/p3157
- Shaw, B., Watson, B., Frauendienst, B., Redecker, A., Jones, T., Hillman, M. (2013) *Children's independent mobility: a comparative study in England and Germany (1971-• 2010)*, Policy Studies Institute, London.
- Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, 95, 1202–1216. doi:10.1037/a0013314

6. Staying Streetwise: The Maintenance of Accurate Judgments of Aggression into Older Age

Manuscript in preparation for publication

Satchell, L., Akehurst, L., Morris, P., & Nee, C. (in preparation) Staying streetwise: the maintenance of accurate judgments of aggression into older age

Abstract

The extant literature has generally demonstrated that individuals can detect the trait aggression of another person with limited information. However there is little research that investigates the life course persistence of aggression detection accuracy. Here, we aimed to explore the accuracy of older adults at detecting potential aggressors. Thirty-nine older adults ($M_{Age}=71.49$ years, $SD_{Age}=7.59$ years) and eighty-seven younger adults ($M_{Age}=20.24$ years, $SD_{Age}=1.74$ years) made intimidation judgments, via video recordings, for nine people (walkers). Our measure of aggression detection accuracy was a comparison of the intimidation judgments made by participants with the walkers' responses to the Buss-Perry Aggression Questionnaire. Both age groups were highly accurate in their recognition of trait aggression and accuracy was maintained into older age, with no difference in accuracy between the older and younger age groups. That said, it was demonstrated that there was more variability in the intimidation ratings made by older adults compared to those made by younger adults. This suggests less consensus in terms of who is intimidating for the older group compared to the younger group. Overall, the adults in this study were highly accurate at detecting trait aggression. Furthermore, there was no meaningful difference in aggression detection accuracy for older adults compared to younger adults. Our finding of a small negative correlation between age and aggression detection accuracy suggests that further research could consider including 'middle aged' adults in order to thoroughly investigate the effect of life course on aggression detection accuracy.

6.1 Introduction

For some time the consensus in the literature was that older adults were more afraid of becoming a victim of crime than younger adults (Clarke & Lewis, 1982; Clemente & Kleinman, 1976; Kennedy & Silverman, 1985; Lewis & Salem, 1986; Moeller, 1989; Ortega & Myles, 1987). However, more recently research has demonstrated that the relationship between age and fear of crime is not so clear, and is mediated by many factors; including the crime type, gender of participants and participants' belief in being able to defend themselves (see Acierno, Rheingold, Resnick & Kilpatrick, 2004; Beaulieu, Dubé, Bergeron & Cousineau, 2007; Jackson, 2009; Kappes, Greve & Hellmers, 2013; Oh & Kim, 2009). Despite the evidence that fear or intimidation is situational, there is little experimental research investigating older adults' detection of potentially dangerous others. It could be that fear of crime is relative to an individual's ability to detect potentially dangerous others with ease. Previous research has demonstrated that simply watching someone walk communicates aspects of dispositional aggression. In particular young adults ($M_{Age} = 19.18$ years) were accurate in their judgments of trait aggression (Chapter 3) and young adults ($M_{Age} = 21.15$ years) were more accurate than teenagers (13-17 years, Chapter 5). The accuracy of older adults, when presented with a similar task, has not yet been investigated. In the current study we investigated whether aggression detection accuracy is maintained into older age.

It is important to make quick, but also accurate, judgments of the danger posed by others. Even with limited interaction, people are generally good at detecting the trait properties of others (Cheek, 1982; Funder, 1980; Letzring, 2015; McCrae, 1982; Vazire & Mehl, 2008) and previous work suggests that accuracy of 'social perceptions' is robust into older age (for an overview see; Freund & Isaacowitz, 2014). Funder's (1995; 1999) Realistic Accuracy Model (RAM) focusses upon the *availability* of *relevant* information

regarding targets' traits, which can then be *detected* by a judge to *utilize* for an accurate personality judgment. This focus on the detection of traits being as much a property of the targets as it is the judges is important, and often overlooked. This is especially important in the context of an approaching person, as the information elicited by a target person changes in quantity and quality as they approach a judge. As the target approaches it will frequently be their gait and their body outline that are *available* (to use the RAM term) to a judge first. Most extant work on the judging of malevolent attributes of others relies on photographs of faces (such as; Geniole, Molnar, Carré & McCormick, 2014; Hehman, Flake & Freeman, 2015; Hehman, Leitner & Gaertner, 2013) and does not include other information potentially relevant to trait aggression which would be available from an approaching person. If older adults can use gait information to detect trait aggression, much like younger adults can (Chapters 3, 4 & 5), then they will be able to make 'approach or avoid' responses when a stranger is walking towards them.

Research has shown that adults are able to accurately judge the dispositional aggression of a target person from simply viewing a photograph of their face (Carré, McCormick & Mondloch, 2009). Boshyan, Zebrowitz, Franklin, McCormick and Carré (2013) have even demonstrated that older adults are as accurate as younger adults in detecting targets' performance in a Point Subtraction Aggression Paradigm (Carré & McCormick, 2008) from viewing brief presentations (3s) of photographs of faces. If older adults are accurate at detecting the traits of others from viewing photographs of faces (in Boshyan et al. the average accuracy value, using fisher's z correlations, of older adults was .20), then it is possible they will also be as accurate, if not more accurate, when judging realistic presentations of a target; a walking person.

For the current experiment we asked participants to judge how 'intimidated' they were by nine individual targets. The targets were presented, sequentially, via video

presentations. We expected the judgments of intimidation would relate to each target's trait aggression (Chapters 3, 4 & 5) and refer to this relationship as *accuracy* (see Funder, 2012). Given that older adults have been shown to be as accurate as younger adults in detecting aggression from photographs of faces, we were also expecting this to be the case in the current experiment. In fact, as our participants were presented with more realistic information than simply the face of a target, we expected our older and younger adults' judgments of aggression would be more accurate than those of previous research that only presented still photographs of faces.

6.2 Method

6.2.1 Participants. Thirty-nine older adults (27 Females, $M_{Age}=71.49$ years, $SD_{Age}=7.59$ years, $Min_{Age}=59$ years, $Max_{Age}=91$ years) were recruited during public lectures delivered to members of an Aging Network research group. These were individuals who engage with academic events hosted by our University. Eighty-seven undergraduate 'younger adults' (73 Females, $M_{Age}=20.24$ years, $SD_{Age}=1.74$ years, $Min_{Age}=18$ years, $Max_{Age}=28$ years) were recruited as a comparison group. The younger adults were recruited and took part in a similar forum to the older adult sample; in an undergraduate lecture.

6.2.2 Materials. This experiment used the same stimulus walkers as a previous study (Chapter 2). We filmed the individuals walking at their preferred speed for 10 seconds on a treadmill. For the purpose of keeping the experiment efficient and retaining the engagement of all participants, a sub-set of nine walkers (from our database of 23) were chosen for use in this study. These nine walkers (5 female, $M_{Age}=20.67$ years, $SD_{Age}=2.40$ years, $Min_{Age}=18$ years, $Max_{Age}=24$ years) were randomly selected using a random number generator.

We used the Physical Aggression subscale from the Buss-Perry Aggression Questionnaire (Buss & Perry, 1992) to capture our walkers' trait aggression. This subscale, when analysed using advised revisions (Bryant & Smith, 2001, see Appendix F), gives a score between 3 and 21 and our nine walkers were reasonably spread in the range of possible scores ($M_{\text{Aggression}}=7.11$, $SD_{\text{Aggression}}=5.09$, $\text{Min}_{\text{Aggression}}=3$, $\text{Max}_{\text{Aggression}}=15$). This subscale was chosen as it is most pertinent to interpersonal aggression. The Buss-Perry Questionnaire is well established in the aggression literature and has been used with various forensic and non-forensic populations to predict historic and laboratory expressions of aggression (see; Archer & Webb, 2006; Bryant & Smith, 2001; Diamond, 2006; Diamond, Wang & Buffington-Vollum, 2005; García-León et al., 2002; O'Connor, Archer & Wu, 2001).

6.2.3 Procedure. Participants took part in groups. The walkers were presented on a screen (measuring approximately 3m x 2.5m) and for each presentation session, the order of presentation of the nine walkers was randomised. After the presentation of each walker, the participants were given as much time as they required to make ratings, on 9-point Likert scales, of how *intimidating-not intimidating*, *friendly-unfriendly* and *masculine-feminine* they perceived each walker to be.

6.2.4 Analyses. We analysed our data in two ways. It is typical of interpersonal perception experiments of this type to calculate the average rating received by the target (in this case each walker) and to correlate this value with the targets' traits to demonstrate whole group accuracy (see *Ratings received by walkers* section of the Results). We included this analysis for the purpose of comparison with other research but this 'nomothetic' approach to analysing judgment data has faced criticism. Such analysis ignores the individual variation in judge ability and falsely increases the size of

correlations (see; Brand & Bradley, 2012; Kolar, Funder & Colvin, 1996; Monin & Oppenheimer, 2005).

As such, we also adopted the recommended ‘idiographic’ analysis (Kolar et al., 1996; that is referred to as the ‘average of correlations’ by Monin & Oppenheimer, 2005), where an accuracy value is calculated for each participant. This is achieved by computing the correlation between judgments of each individual judge and the targets’ properties, thus producing a value between 1 (accurate; a walker rated as more intimidating is more aggressive) and -1 (inaccurate; a walker rated as more intimidating is less aggressive) and the size of each value denotes the strength of agreement (0 being random performance, ± 1 being perfect [in]accuracy). The idiographic analysis allows us to report the distribution of judge variability, as well as to test for the effects of age on accuracy (see *Accuracy of participants* section of the Results).

6.3 Results

6.3.1 Ratings received by walkers. The average intimidation ratings the walkers received positively correlated with the walkers’ trait aggression, with notably large effects (younger adults, $r(9)=.85$, 95% CI [.37, .99], $p=.004$; older adults, $r(9)=.82$, 95% CI [.48, .96], $p=.007$). Thus, both groups demonstrated very strong accuracy at detecting trait aggression through intimidation ratings (there was no meaningful difference between the two groups’ accuracy correlations; $z=.17$, $p=.865$.) Further, we found that the average rating received by the walkers from the older adults and younger adults strongly correlated ($r(9)=.79$, 95% CI [.46, 1.00], $p=.012$) suggesting similar allocation of intimidation ratings. It is interesting to note that there were no sex differences in the aggression of the walkers and as such the judges did not appear more accurate by labelling male targets as more aggressive ($U(N=9)= 5.50$, $p=.268$, see figure 6.1).

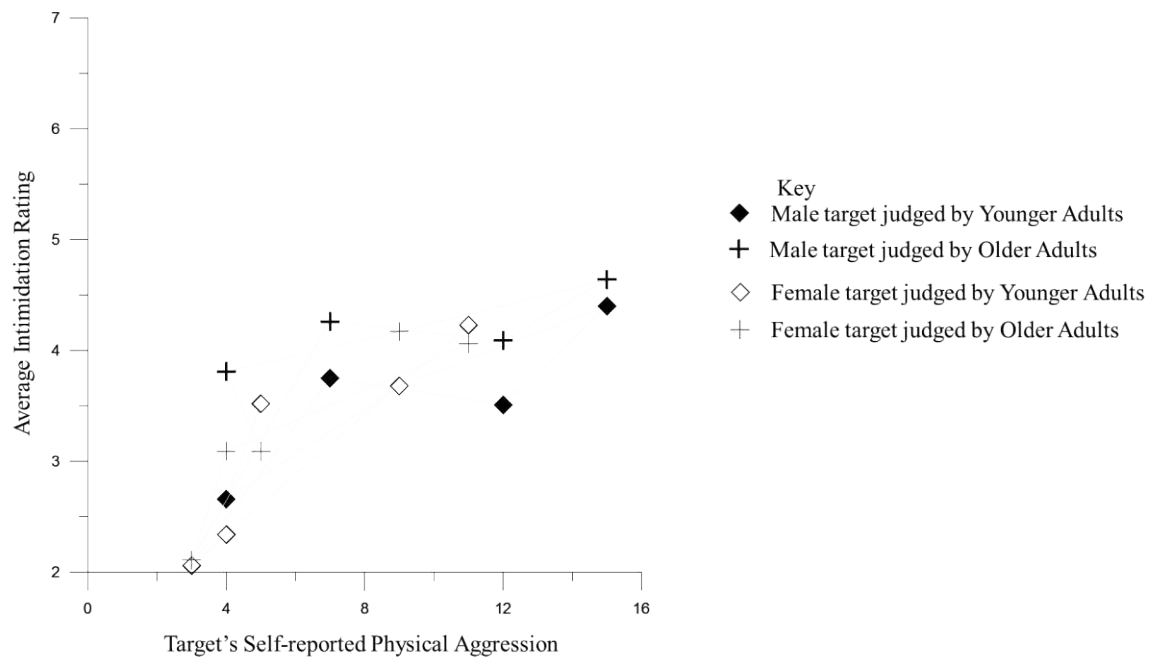


Figure 6.1 The relationship between average judgment of intimidation received by the walkers and the walker's trait aggression. Results are presented by age group and sex of walker.

In all measures in psychology it is typical to report both the mean and the standard deviation of a distribution. It is somewhat surprising that similar research does not report the variation in ratings received by the targets. Variance in judgments demonstrates more guess work or a lack of consensus about the matter being judged. By calculating the standard deviation (σ) of intimidation ratings received by each walker, we were able to test for differences in spread of intimidation judgments between the age groups. We found that the variation in intimidation ratings differed between age groups ($t(8)=2.70, p=.027, d=.76$) with the older adults being more varied in their ratings ($M_{\sigma}=2.02, SD_{\sigma}=.28$) than the younger adults ($M_{\sigma}=1.78, SD_{\sigma}=.35$). The implication of these results is that whilst, on average, the groups performed similarly, there was less consistency in ratings for the older adults compared to the younger adults.

6.3.2 Accuracy of participants. Given the variation in ratings demonstrated above, it was important to consider the distribution of judge accuracy (see figure 6.2). Thus we report the distribution of individual accuracy correlations (r , where $r=1$ is perfect accuracy, $r=-1$ perfect inaccuracy and $r=0$ random responding). There was no real difference in accuracy⁸ between the older adults ($M_r=.35$, $SD_r=.26$, $Min_r=-.24$, $Max_r=.91$, Skewness= -.24, Kurtosis= -.35) and younger adults ($M_r=.43$, $SD_r=.27$, $Min_r=-.48$, $Max_r=.92$, Skewness= -.77, Kurtosis= .84) ($t(124^9)=1.61$, $p=.110$, $d=.31$). It is important to note that accuracy was generally good across both age groups.

There was a limited distribution of age, with a notable lack of participants between the ages of 28 and 59 years, so correlations between age and accuracy should be treated with caution. Using a Spearman's ranked correlation (to prevent too much influence of this age gap on our correlations) we found a small but significant negative correlation between age and accuracy ($r_s(126)=-.20$, $p=.029$, and this effect was weaker in a Pearson's correlation; $r(126)=-.15$, 95% CI [-.32, .01], $p=.092$).

It is notable that some participants were extremely accurate (6 participants achieved $r>.80$, 5 of whom were in the younger adults group). In some cases there were very small differences in the self-report measures of the walkers (in some cases a 1 point difference in their self-reported aggression), yet some participants were able to detect these subtle differences.

⁸ We opted not to use Fisher's Z transformation of our correlational data for our analysis due to concerns about artificially increasing effect sizes after 'normalising' the distribution of r values. If transformed the individuals *appeared* to be more accurate when reporting means in both the younger adults ($M_{z(r)}=.51$, $SD_{z(r)}=.37$) and older adults ($M_{z(r)}=.41$, $SD_{z(r)}=.35$).

⁹ Variance in individuals' accuracy between groups were relatively equal despite differences in group size ($SD_{Older\ Adults}=.26$, $SD_{Younger\ Adults}=.27$, Levene's $W=.00$, $p=.959$)

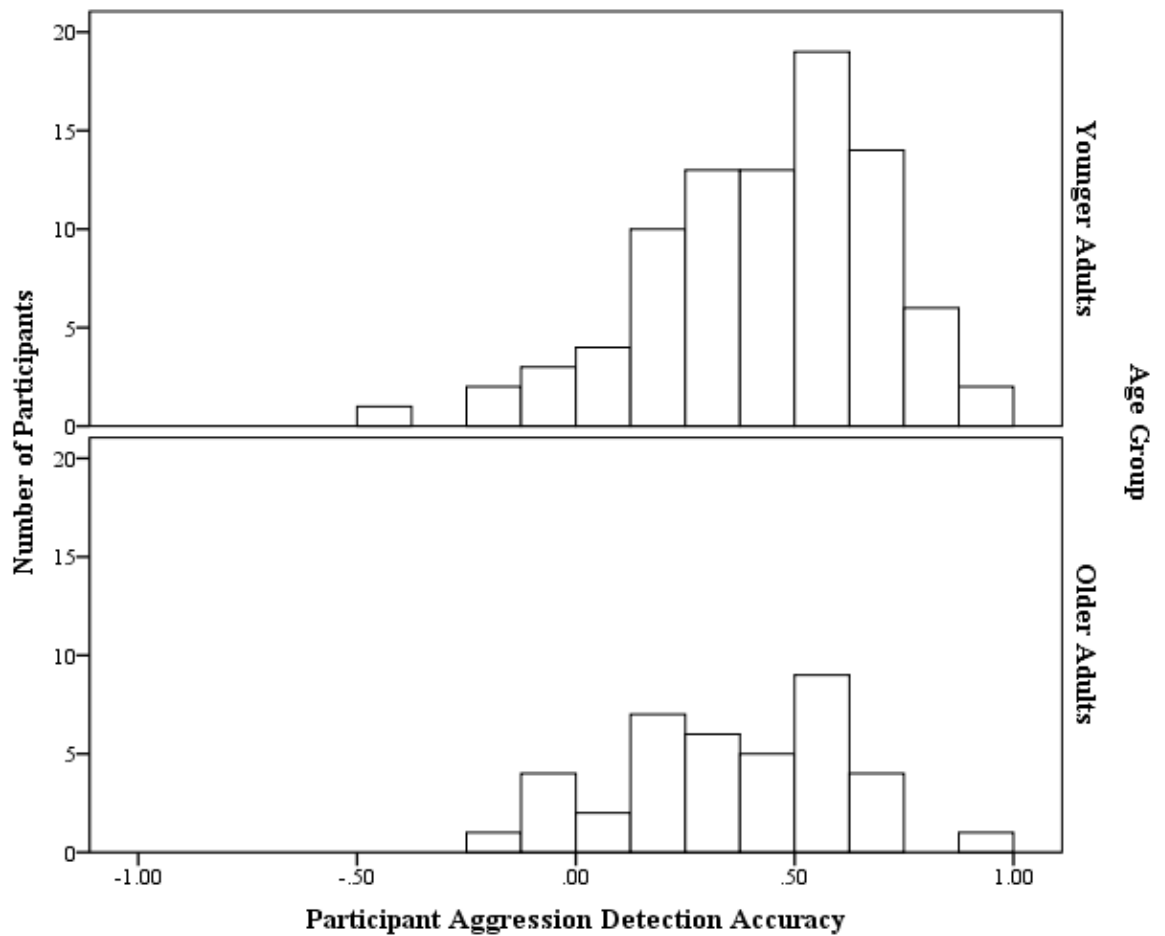


Figure 6.2. The distribution of participant accuracy (individual correlations) for each age group.

6.4 Discussion

In this study we found that older adults were as accurate as younger adults at detecting the trait aggression of our targets. We found that the ratings received by the targets from both the older and younger adult groups correlated strongly with each other and with the trait aggression of the targets. In fact, our measures showed that both age groups were reasonably good at indexing trait aggression through intimidation ratings, with some participants being notably good at using their intimidation ratings to reflect even small differences in trait aggression for the targets. There was some evidence of a negative correlation between age and accuracy, but the sample lacked participants between the ages

of 30 and 55 years so this finding should be treated with caution. Our categorical analyses (younger adults vs older adults) demonstrated no difference between the groups in terms of accuracy. The only clear difference between the groups was in the spread of ratings received by the targets. The younger participants showed more consensus in intimidation ratings than the older adults. Whilst this did not impact upon accuracy (it is apparent that the discrimination between more and less dangerous targets is constant between groups), it may show that there is less specificity in what intimidation means to the older adults.

The fact that intimidation ratings, which do not explicitly target the concept of aggression, relate to the ordering of the walkers' aggression is interesting and somewhat surprising. Given that some of the self-report aggression ratings made by the walkers were very similar, it is unexpected that some participants achieved such high accuracy. It is also important that, whilst the relationship is implicit, intimidation is not a direct reference to aggression. Boshyan et al. (2013) asked participants to form explicit aggression judgments about photographs of targets and they told their participants about the paradigm through which the aggression values for each target had been attained (a computer task called the Point Subtraction Aggression Paradigm). In the current study, we did not inform participants of our intentions to test them for accuracy thus allowing them to make a general judgment of intimidation. Our participants did not know that their judgments would be compared with the self-reported aggression of the targets. Interestingly, the current findings replicated the findings of Boshyan et al. (2013) and found evidence of strong accuracy in the older population¹⁰.

The inconsistencies in the relationship between age and fear of crime could be due to individual differences. As previous research shows, the idiosyncrasies of respondents,

¹⁰ For the older participants in Boshyan et al.'s study, they reported a Fisher's Z transformed accuracy value of .20. For the current study the Fisher's Z transformation of accuracy for the adults was .41.

context and crime type all have an effect on fear of crime judgments in older adults (Acierno, et al., 2004; Beaulieu, et al., 2007; Jackson, 2009; Kappes, et al., 2013; Oh & Kim, 2009). It is possible that older adults who are reasonably accurate at detecting potential dangers may well be those adults who are less fearful of crime. Asymmetry between feelings of intimidation and genuine danger could lead to psychological and physical health consequences. A growing body of literature has demonstrated that how much an individual fears crime in their local area effects the time they spend walking (Foster, Knuiman, Hooper, Christian & Giles-Corti, 2014), cycling (Kramer, Maas, Wingen & Kunst, 2013) and how much time older adults spend engaging in healthy activity programs (Dawson, Hillsdon, Boller & Foster, 2007). Perhaps dissemination of research findings that find that older people's judgments of dangerous others are largely accurate may provide individuals with encouragement to engage with their environment. Equally, it could be the case that those who feel more intimidated in general are those with poor accuracy. Future research might usefully evaluate the possibility of improving accuracy through training.

We were interested in demonstrating that older adults can be accurate at judging danger via realistic presentations of targets. As our stimuli (videos) were time-consuming to present, only nine randomly-selected targets (from a larger sample) were presented to participants. Future work could consider more targets and perhaps, rather than presenting targets one at a time, they could be presented in groups (something that many of the older participants self-reported as being important for their everyday judgments of intimidation). It would also be of interest to investigate the detection of non-physical risk. As Acierno et al. (2004) note, perceived risk of property crime is more relevant for older adults than a physical crime, so perhaps a test of accuracy in detecting a target person's propensity to

commit acquisitive crime could be of interest (for example, in the context of doorstep crime).

6.4.1 Conclusion. In sum, older adults were as accurate as younger adults at detecting the trait aggression of walking targets using ratings of intimidation. Previous research has suggested that experience is essential for the acquisition of accurate intimidation judgments (Chapters 3 & 4) and so it is reasonable to suggest that accuracy does not decline as people attain experience throughout their lives. Whilst the current findings need to be replicated, with more participants, more targets and a greater age range of participants and targets, they convey a positive message for older adults in that their gut reactions pertaining to the threat posed by approaching strangers are likely to be accurate.

References

- Acierno, R., Rheingold, A. A., Resnick, H. S., & Kilpatrick, D. G. (2004). Predictors of fear of crime in older adults. *Journal of Anxiety Disorders*, 18(3), 385–396. doi:10.1016/s0887-6185(03)00012-4
- Archer, J., & Webb, I. A. (2006). The relation between scores on the Buss–Perry Aggression Questionnaire and aggressive acts, impulsiveness, competitiveness, dominance, and sexual jealousy. *Aggr. Behav.*, 32, 464–473. doi:10.1002/ab.20146
- Beaulieu, M., Dubé, M., Bergeron, C., & Cousineau, M.-M. (2007). Are elderly men worried about crime? *Journal of Aging Studies*, 21(4), 336–346. doi:10.1016/j.jaging.2007.05.001
- Boshyan, J., Zebrowitz, L.A., Franklin, R.G., McCormick, C.M., & Carré, J.M. (2013). Age similarities in recognizing threat from faces and diagnostic cues. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 69(5), 710–718, doi:10.1093/geronb/gbt054.
- Boulton, M. J., & Smith, P. K. (1994). Bully/victim problems in middle school children: Stability, self-perceived competence, peer perceptions and peer acceptance. *British journal of developmental psychology*, 12, 315-330. doi:10.1111/j.2044-835x.1994.tb00637.x
- Brand, A., & Bradley, M. T. (2012). More voodoo correlations: when average-based measures inflate correlations. *The Journal of general psychology*, 139(4), 260-272. doi:10.1080/00221309.2012.703711
- Bryant, F. B., & Smith, B. D. (2001). Refining the Architecture of Aggression: A Measurement Model for the Buss–Perry Aggression Questionnaire. *Journal of Research in Personality*, 35, 138–167. doi:10.1006/jrpe.2000.2302

- Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology*, 63, 452–459. doi:10.1037/0022-3514.63.3.452
- Carré, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behavior in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 275, 2651–2656. doi:10.1098/rspb.2008.0873
- Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial Structure Is a Reliable Cue of Aggressive Behavior. *Psychological Science*, 20, 1194–1198. doi:10.1111/j.1467-9280.2009.02423.x
- Cheek, J. M. (1982). Aggregation, moderator variables, and the validity of personality tests: A peer-rating study. *Journal of Personality and Social Psychology*, 43, 1254–1269. doi:10.1037/0022-3514.43.6.1254
- Clarke, A. H., & Lewis, M. J. (1982). Fear of crime among the elderly. *British Journal of Criminology*, 22, 49–62.
- Clemente, F., & Kleiman, M. B. (1977). Fear of crime in the United States: A multivariate analysis. *Social Forces*, 56, 519–531. doi: 10.1093/sf/56.2.519
- Costall, A. (1995). Socializing affordances. *Theory & Psychology*, 5(4), 467–481. doi:10.1177/0959354395054001
- Dawson, J., Hillsdon, M., Boller, I., & Foster, C. (2007). Perceived barriers to walking in the neighbourhood environment and change in physical activity levels over 12 months. *British Journal of Sports Medicine*, 41(9), 562–568. doi:10.1136/bjsm.2006.033340
- Diamond, P. M. (2006). The Short-Form Buss-Perry Aggression Questionnaire (BPAQ-SF): A Validation Study With Federal Offenders. *Assessment*, 13, 227–240. doi:10.1177/1073191106287666

- Diamond, P. M., Wang, E. W., & Buffington-Vollum, J. (2005). Factor structure of the Buss-Perry Aggression Questionnaire (BPAQ) with mentally ill male prisoners. *Criminal Justice and Behavior*, 32, 546-564.
doi:10.1177/0093854805278416
- Fessler, D. M. T., Holbrook, C., Pollack, J. S., & Hahn-Holbrook, J. (2014). Stranger danger: Parenthood increases the envisioned bodily formidability of menacing men. *Evolution and Human Behavior*, 35, 109–117.
doi:10.1016/j.evolhumbehav.2013.11.004
- Foster, S., Knuiman, M., Hooper, P., Christian, H., & Giles-Corti, B. (2014a). Do changes in residents' fear of crime impact their walking? Longitudinal results from reside. *Preventive medicine*, 62, 161-166. doi:10.1016/j.ypmed.2014.02.011
- Freund, A. M., & Isaacowitz, D. M. (2014). Aging and social perception: So far, more similarities than differences. *Psychology and Aging*, 29(3), 451–453.
doi:10.1037/a0037555
- Funder, D. C. (Ed.). (1999). *Personality judgment: A realistic approach to person perception*. Academic Press.
- Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, 21, 177-182. doi:10.1177/0963721412445309
- García-León, A., Reyes, G. A., Vila, J., Pérez, N., Robles, H., & Ramos, M. M. (2002). The Aggression Questionnaire: A validation study in student samples. *The Spanish Journal of Psychology*, 5, 45-53. doi:10.1017/s1138741600005825
- Geniole, S. N., Molnar, D. S., Carré, J. M., & McCormick, C. M. (2014). The facial width-to-height ratio shares stronger links with judgments of aggression than with judgments of trustworthiness. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1526–1541. doi:10.1037/a0036732

- Golinkoff, R. M., Chung, H. L., Hirsh-Pasek, K., Liu, J., Bertenthal, B. I., Brand, R., ... Hennon, E. (2002). Young children can extend motion verbs to point-light displays. *Developmental Psychology*, 38, 604–614. doi:10.1037/0012-1649.38.4.604
- Helman, E., Flake, J. K., & Freeman, J. B. (2015). Static and Dynamic Facial Cues Differentially Affect the Consistency of Social Evaluations. *Personality and Social Psychology Bulletin*, 41, 1123-1134. doi:10.1177/0146167215591495
- Helman, E., Leitner, J. B., & Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology*, 49(4), 747-754. doi:10.1016/j.jesp.2013.02.015
- Hirschmüller, S., Egloff, B., Schmukle, S. C., Nestler, S., & Back, M. D. (2015). Accurate judgments of neuroticism at zero acquaintance: A question of relevance. *Journal of personality*, 83(2), 221-228. doi:10.1111/jopy.12097
- Jackson, J. (2009). A psychological perspective on vulnerability in the fear of crime. *Psychology, Crime & Law*, 15(4), 365–390. doi:10.1080/10683160802275797
- Kappes, C., Greve, W., & Hellmers, S. (2013). Fear of crime in old age: precautionary behaviour and its relation to situational fear. *European Journal of Ageing*, 10, 111–125. doi:10.1007/s10433-012-0255-3
- Kennedy, L. W., & Silverman, R. A. (1985). Significant others and fear of crime among the elderly. *International Journal on Aging and Human Development*, 20, 241–256.
- Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of personality*, 64, 311-338. doi:10.1111/j.1467-6494.1996.tb00513.x
- Kramer, D., Maas, J., Wingen, M., & Kunst, A. E. (2013). Neighbourhood safety and leisure-time physical activity among Dutch adults: a multilevel perspective.

- International Journal of Behavioral Nutrition and Physical Activity*, 10.
doi:10.1186/1479-5868-10-11
- Letzring, T. D. (2015). Observer judgmental accuracy of personality: Benefits related to being a good (normative) judge. *Journal of Research in Personality*, 54, 51–60.
doi:10.1016/j.jrp.2014.05.001
- McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology*, 43, 293–303.
doi:10.1037/0022-3514.43.2.293
- McNaughton, N., & Corr, P. J. (2004). A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance. *Neuroscience & Biobehavioral Reviews*, 28, 285-305. doi:10.1016/j.neubiorev.2004.03.005
- Moeller, G. L. (1989). Fear of criminal victimization: the effect of neighborhood racial composition. *Sociological Inquiry*, 59, 208–221. doi: 10.1111/j.1475-682X.1989.tb00101.x
- Monin, B., & Oppenheimer, D. M. (2005). Correlated averages vs. averaged correlations: Demonstrating the warm glow heuristic beyond aggregation. *Social Cognition*, 23(3), 257-278. doi:10.1521/soco.2005.23.3.257
- Moran, E., Warden, D., Macleod, L., Mayes, G., & Gillies, J. (1997). Stranger Danger: What do Children Know?. *Child Abuse Review*, 6, 11-23. doi: 10.1002/(SICI)1099-0852(199703)6:1<11::AID-CAR286>3.0.CO;2-G
- Mynard, H., & Joseph, S. (1997). Bully/victim problems and their association with Eysenck's personality dimensions in 8 to 13 year olds. *British Journal of Educational Psychology*, 67(1), 51-54. doi: 10.1111/j.2044-8279.1997.tb01226.x

- O'Connor, D. B., Archer, J., & Wu, F. W. C. (2001). Measuring aggression: Self-reports, partner reports, and responses to provoking scenarios. *Aggressive Behaviour*, 27, 79–101. doi:10.1002/ab.2
- Oh, J.-H., & Kim, S. (2009). Aging, neighborhood attachment, and fear of crime: testing reciprocal effects. *J. Community Psychol.*, 37(1), 21–40. doi:10.1002/jcop.20269
- Ortega, S. L., & Myles, J. L. (1987). Race and gender effects on the fear of crime: an interactive model with age. *Criminology*, 25, 133–152. doi: 10.1111/j.1745-9125.1987.tb00792.x
- Pavlova, M., Krägeloh-Mann, I., Sokolov, A., & Birbaumer, N. (2001). Recognition of point-light biological motion displays by young children. *Perception*, 30, 925–933. doi:10.1068/p3157
- Shaw, B., Watson, B., Frauendienst, B., Redecker, A., Jones, T., Hillman, M. (2013) *Children's independent mobility: a comparative study in England and Germany (1971-• 2010)*, Policy Studies Institute, London.
- Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, 95, 1202–1216. doi:10.1037/a0013314

7. Describing Individual Differences in Targets' Salience and Judges' Detection of Trait Aggression

Based on the manuscript under review with the Journal of Personality

Satchell, L. & Morris, P. (under review) *Describing individual differences in targets' salience and judges' detection of trait aggression*

Abstract

Many studies of interpersonal trait identification use analyses which ignore potential variance in judges and targets of traits. Research in recognising a target's traits should move to researching individual differences in judge performance and the salience of target's traits. In the context of a trait aggression detection experiment, we compare and contrast standard analysis techniques with an 'idiographic' approach that takes into account individual differences in both judge performance and availability of targets' aggression. 58 undergraduate students judged the threat of 22 videos of target individuals who were walking in an experiment where judges were oriented towards or away from the stimuli. Results show that participants' judgments of threat did relate to the trait aggression of the targets. This result was demonstrated using the popular approach of correlating the aggregation of judge ratings with the targets' trait scores. We also use correlations for individual judges to demonstrate the judge ability is normally distributed in this study. By using a novel analysis we then demonstrate that one target had disproportionately salient traits. As an outlier, this target was greatly increasing the reported accuracy of the judges. In sum, to thoroughly engage with the individual differences in trait recognition, detail on the distribution of judge and target properties should be reported. Some individual's traits are very available to judges and the influence of individual targets should be considered when analysing data.

7.1 Introduction

Individual differences psychologists have studied the ability of participants to detect individual differences in the psychology of others (Albright, Kenny & Malloy, 1988; Funder, 1980; 2012; McCrae, 1982). As Funder (1999) proposes in his Realistic Accuracy Model (RAM), the identification of another person's traits is dependent on both the judge's ability to recognise traits and the salience of the trait being detected in the targets. The RAM proposes that accurate judgments of another person's traits are dependent on four criteria of a judge-target dyad being met. A target's features or behaviour must relate to the trait being judged, some attribute of the target must be *relevant* to the judgment. Those attributes of the target must then be *available* (observable) for the judge to *detect*. If the judge does detect the relevant attributes of the target, they still must correctly *utilize* the information to make an accurate judgment. It is, perhaps surprising, that the typical approach to analysing trait judgment experiments ignores individual variation in both judges and targets, especially given that it is reasonable to assume that there could be individual variation (in judges and targets) for each stage of the RAM. This paper revisits existing methods of analysing judge variation and promotes analysis of target variation.

Kolar, Funder and Colvin (1996) observe that using 'nomothetic' approaches to analysing trait recognition data could lead to problems in reporting true judge accuracy. The nomothetic analysis of trait identification data uses the average rating judgment made by a sample of judges as a target property, to be correlated with that target's traits. So for example, a group of ten judges make ratings of one stimulus person. The judges' ratings are then averaged to produce one rating value for one stimulus. This process is repeated for all the stimuli. Then this average rating received by the target is correlated with those targets' traits. This, in effect, measures how the judges, as a whole, can predict the traits of the targets. However, this kind of nomothetic analysis has been demonstrated to falsely

increase the effect size of social perception correlations due to the removal of the individual differences in judge perception (Monin & Oppenheimer, 2005).

Despite clear concerns with nomothetic analyses, trait identification studies are still frequently analysed in this way (a few examples are; Back, Schmukle, & Egloff, 2008; Carré, McCormick, & Mondloch, 2009; Coetzee, Chen, Perrett, & Stephen, 2010; Feinberg, DeBruine, Jones, & Perrett, 2008; Gillath, Bahns, Ge, & Crandall, 2012; Hall, Pennington, & Lueders, 2013; Harms, Han, & Chen, 2012; Harrison, Shortall, Dispenza, & Gallup, 2011; Holleran, & Mehl, 2008; Holtzman, 2011; Nestler, Egloff, Küfner, & Back, 2012; Rule, & Ambady, 2008; Savani, Stephens, & Markus, 2011). Brand and Bradley (2012) summarised the problems with correlations based on aggregated judgements, likening nomothetic analyses to the ‘voodoo correlations’¹¹ in neuropsychology research. However, average judge responses are still the most popular format of reporting in interpersonal accuracy experiments (a few more recent examples are; Doll, et al., 2014; Fong, & Mar, 2015; Fruhen, Watkins, & Jones, 2015; Little, Jones, DeBruine, & Dunbar, 2013; Sim, Saperia, Brown, & Bernieri, 2015).

There has been some work to move away from using aggregated judge data. Kolar et al. (1996) propose analysing judgment accuracy on a judge by judge basis, in what they termed ‘idiographic’ judge analyses (a few examples are; Hirschmüller, Egloff, Schmukle, Nestler, & Back, 2015; Lippa, & Dietz, 2000; Naumann, Vazire, Rentfrow, & Gosling, 2009). These are correlations of each judge’s accuracy that can be reported as a

¹¹ Brand and Bradley (2012) adopt the term ‘voodoo correlations’ to define any misrepresentative correlation based on aggregate data. They define the origins of the term ‘voodoo correlations’ as the following;

“a correlation estimate can be inflated or deflated by non-representative sampling of the experimental stimuli (Fielder, 2011). In other words, researchers may, inadvertently or advertently, select stimuli to maximize the correlation between two measures rather than try and select stimuli that are representative of the stimuli population. High-profile examples of distorted correlations are from the inflated so called “Voodoo” correlations in social neuroscience. These result from non-independent sampling of volume elements or voxels in fMRI studies (Vul, Harris, Winkielman, & Pashler, 2009) where reported correlations were derived from only the voxels that were strongly correlated with the behavioral measure.” (p.261)

distribution (Monin & Oppenheimer, 2005). These analyses get closer to understanding the individual differences involved in judges' abilities to recognise traits and are reported throughout the thesis (Chapters 4, 5 & 6). It is essential to understand the variation in judges' ability to identify traits, when many personality judgment studies draw conclusions from only a few judges, with the assumption that identification of traits is consistent across judges (e.g. using 6-9 judges of targets' traits, Holleran & Mehl, 2008).

Whilst there is a move to include analysis of judge *detection* and *utilization* of information (to use RAM terminology, Funder's, 1995; 1999), Kolar et al.'s (1996) idiographic judge analysis still assumes that target traits are normally distributed in salience of traits. That is to say all targets have features that are equally *relevant* to their traits and that these features are equally *available*. If one target's traits are more or less salient than the others in any of the previous studies, the (idiographic judge and nomothetic) correlations, could be artificially bolstered or diminished, leading to type I or type II error respectively. In this paper we advocate the reporting of 'idiographic target' trait salience as well as the idiographic judge accuracy.

We use data from an exploratory study into aggression detection through threat ratings to demonstrate the effects of the different analysis styles. The experiment involved participants making a judgment of the threat posed walking targets that the judges were either facing or had to turn to face. The experiment was intended to explore differences in accuracy of threat judgments between the two conditions and was based on research that demonstrated that disliking or avoidance of others can be manifest by turning away from that person (McCall, Blascovich, Young & Persky, 2009; McCall & Singer, 2015). It was considered that, through the embodiment of being oriented towards or away from a target of threat ratings, threat judgments could become less accurate. In the current experiment there was no difference in accuracy between being oriented towards or away from target

presentation. Instead, we dedicate this data to advocating individual difference reporting styles in similar future research.

7.2 Method

7.2.1 Participants. 58 undergraduate students (Female= 46, $M_{\text{Age}} = 18.47$, $SD_{\text{Age}} = .88$) took part in the experiment for a course credit. Participants were informed that they would be taking part in an experiment on interpersonal perception called “They’re Behind You?”

7.2.2 Stimulus Materials. The targets in this experiment were 22 individuals (Female = 11, $M_{\text{Age}} = 20.50$, $SD_{\text{Age}} = 2.04$) who were recorded walking on a treadmill at their chosen speed. The targets were oriented towards the camera and a video camera recorded 5 seconds of uninterrupted gait. Targets wore standardised clothes; for male targets a white t shirt and blue shorts and for female targets a grey vest top and black leggings.

As a measure of trait aggression, targets completed the Buss-Perry Aggression Questionnaire (Buss & Perry, 1992, analysed using revisions suggested by Bryant & Smith, 2001). We use the physical aggression subscale for analysis as it is most relevant to detecting genuine inter-human threat. As per Bryant and Smith’s (2001, see Appendix F) revisions participants could score between 3 and 21 for the physical aggression measure. In our sample of targets there was a reasonable spread of scores ($M_{\text{Aggression}} = 7.46$, $SD_{\text{Aggression}} = 4.78$, $Min_{\text{Aggression}} = 3$, $Max_{\text{Aggression}} = 19$).

7.2.3 Procedure. Participants were randomly allocated to a ‘towards’ or ‘away’ condition. Participants stood 4m away from a presentation screen and judge how

*Threatening-Non Threatening, Masculine-Feminine and Attractive-Unattractive*¹² they found the targets on 9 point Likert scales (on a clipboard). Participants were stood in ‘footprints’, with their feet between wooden blocks (15cm apart so as to act as a guide but not be restrictive) to keep them either facing directly ‘towards’ the presentation of the targets or 120° ‘away’ from the screen. Participants in the away condition were instructed to keep their body aligned with their feet and look over their shoulder at the stimuli, but this was not enforced. Participants all self-reported no history of neck or back problems. Participants completed the experiment alone, but were observed by an experimenter to check that they remained in the footprints. The 22 targets were presented in a random order. Participants had as much time as they needed to complete their ratings before moving onto the next video, controlling the experiment with a presentation remote.

7.3 Analysis and Results

7.3.1 The nomothetic approach; using the average rating received by the targets. The traditional approach to analysing the relationship between two properties of a target (be they perceptual or trait) requires the data analyst to aggregate the data across judges to produce a single data point per target (). For example, we could analyse in the relationship between the average judge ratings of target masculinity and aggregated judgements of threat; how do judgments of masculinity and threat relate? Using similar analyses, we could test for the relationship between average threat rating received by the target and the target’s trait aggression. This is the correlation between the average rating received by the targets and the traits of the targets (τ_k); $\rho(\bar{r}_k, \tau_k)$.

In an experiment with multiple conditions, we can use this analysis as a measure of agreement between two conditions. In this experiment the average threat rating received by targets in towards condition ($M_{\text{Threat Towards}} = 2.71$, $SD_{\text{Threat Towards}} = .77$, Min_{Threat}

¹² This was included as a distractor variable, it is not a variable analysed here.

Towards=1.48, $Max_{Threat\ Towards}=4.97$) were highly similar to those in the away condition ($M_{Threat\ Away}=3.01$, $SD_{Threat\ Away}=.89$, $Min_{Threat\ Away}=1.86$, $Max_{Threat\ Towards}=5.76$, $r(22)=.92$, $95\% CI [.81, .96]$ $p<.001$). The implication of this agreement between conditions is; there is no difference between the groups. In fact the average threat ratings received by the targets from the sample as a whole ($M_{Threat\ Altogether}=2.86$, $SD_{Threat\ Altogether}=.81$, $Min_{Threat\ Altogether}=1.74$, $Max_{Threat\ Altogether}=5.36$) relate so strongly to the ratings received by the targets in the towards ($r(22)=.98$, $95\% CI [.95, .99]$ $p<.001$) and away ($r(22)=.98$, $95\% CI [.96, .99]$, $p<.001$) condition, we can analyse the sample as a whole (without dividing by conditions).

With the target ratings received from the sample as a whole, we find that the average threat rating received by the targets positively correlated with the average masculinity rating received by the targets ($M_{Average\ Masculinity}=4.47$, $SD_{Average\ Masculinity}=1.93$, $Min_{Average\ Masculinity}=1.84$, $Max_{Average\ Masculinity}=7.52$, $r(22)=.77$, $95\% CI [.51, .92]$, $p<.001$, see figure 7.1A); a more masculine target was also seen as more threatening. The average threat ratings received by the target also positively correlated with to the trait aggression of the targets ($r(22)=.43$, $95\% CI [-.02, .78]$, $p=.043$, see figure 7.1B), showing that threat judgments reflect the aggression tendencies of the target. Interestingly, despite the relationship between threat and masculinity ratings, there was a very small relationship between the average masculinity rating received by the targets and the targets' trait aggression ($r(22)=.18$, $95\% CI [-.27, .62]$, $p=.430$).

There is an issue with the use of an aggregation of judge ratings. When describing a distribution of data in any psychology journal, it is expected that *both* the mean and *standard deviation* of a sample is reported. This reduction of 58 judges' ratings down to a single point for analysis ignores the variability in both judge ability and ratings received by targets. Inconsistency in how that target was rated by the sample of judges can be

demonstrated in the standard deviation in threat ratings received by each target ($M_{\text{Variance}}=1.78$, $SD_{\text{Variance}}=.39$, $Min_{\text{Variance}}=.98$, $Max_{\text{Variance}}=2.28$). This measure of variance in ratings received by the target not only better describes the data but also can inform novel findings (such as the development of rating cohesion with age, Chapter 5).

When using the average judge rating as a dependent measure, the analysis assumes that the judges' ability to detect traits is normally distributed, in other words that the mean is a meaningful reflection of that sample (and as such skewness \approx 0 and kurtosis \approx 0). Whilst we do not doubt that in many cases it is possible that a normal distribution is present in performance, we do note that it is rarely reported.

7.3.2 Idiographic descriptions of judge accuracy; individual correlations. It is possible report to the distribution of judge accuracy using Pearson's correlations (such as; Hirschmüller, Egloff, Schmukle, Nestler & Back, 2015; Kolar, Funder & Colvin, 1996;). Each judge rates each target in our paradigm and so it is possible to look at the correlation between the ratings made by the judge of each target (j_k) and the traits of the target being rated (τ_k); $\rho(j_k, \tau_k)$. This gives an individual correlation for each judge, thereby treating 'accuracy' as a judge property ($M_{\text{Accuracy}}=.24$, $SD_{\text{Accuracy}}=.20$, $Min_{\text{Accuracy}}=-.29$, $Max_{\text{Accuracy}}=.57$). This can be interpreted like any correlation, with a score of 1 demonstrating high accuracy (e.g. more threatening targets are more aggressive), a score of -1 demonstrating high inaccuracy (e.g. more threatening targets are less aggressive) and a score of 0 demonstrating no relationship between threat ratings and trait aggression (random performance).

There are multiple questions that can now be answered with this form of analysis. We can demonstrate that there were no differences between the orientation conditions in accuracy ($t(56)=.19$, $p=.848$, $d=.05$) and we can demonstrate that there was no relationship between judge accuracy and judge age ($r(58)=.09$, 95% CI [-.34, .16], $p=.497$). We can

also demonstrate that judge sex had a moderate effect on accuracy ($t(29.68)=4.19$, $p<.001$, $d=1.15$, albeit with unbalanced groups; Female=46, Levene's $W=4.18$, $p=.046$) with males ($M_{\text{Accuracy}}=.39$, $SD_{\text{Accuracy}}=.12$) having higher accuracy than females ($M_{\text{Accuracy}}=.20$, $SD_{\text{Accuracy}}=.20$). This experiment in this paper was not designed to nor was it expected to elicit sex or age differences (we used an undergraduate sample so our participants were overwhelmingly female and 18-22 years old); we simply report tests to demonstrate the utility of this measure of judge accuracy.

One of the assumptions made when using the nomothetic, average rating received by the target, analysis (as is done in section 3.1) is that the performance of the judges is normally distributed. The data from the accuracy of our whole sample were normally distributed (Kurtosis= -.34, Skewness= -.47, see figure 7.2B), it has been recommended that skewness should be no greater than .496 with a sample of 60 participants (Doane & Steward, 2011). However, as Blanca, Arnau, López-Montiel, Bono and Bendayan (2013) note; in real samples of social science 'normal' distributions rarely occur (only 5.5% of 693 distributions were 'normal' as per typical cut-offs for skewness and kurtosis). It is therefore reasonable to have caution when assuming 'normality' of this type of data.

Monin and Oppenheimer (2005) advocate the reporting of the *average of correlations* as well as the *correlation of averages* when dealing with two social perception variables. In our case, it would be of interest to investigate the individual variation in the relationship between ratings of threat and ratings of masculinity. This measure does reveal some surprising features in the data not apparent with a single aggregated value for the judges ($M_{\text{Masculine-Threat}}=.49$, $SD_{\text{Masculine-Threat}}=.24$, $Min_{\text{Masculine-Threat}}=.03$, $Max_{\text{Masculine-Threat}}=.84$, Kurtosis= -1.01, Skewness= -.45, see figure 7.2A). Whilst the sample in general considered masculinity an index of threat, this (somewhat peaky) distribution shows that 13.80% of participants demonstrate an arguably negligible ($r<.20$) relationship between

these two perception ratings. These individuals understand the relationship between masculinity and threat differently to the others in the sample. Just as Monin and Oppenheimer (2005) report, we no difference in the essential message of the results but we gain information from the surprising variance in this relationship. Whilst this analysis may not show differences in the concept being investigated to the previous correlation of averages, as individual differences psychologists, judge idiosyncrasy should be of interest when describing our data.

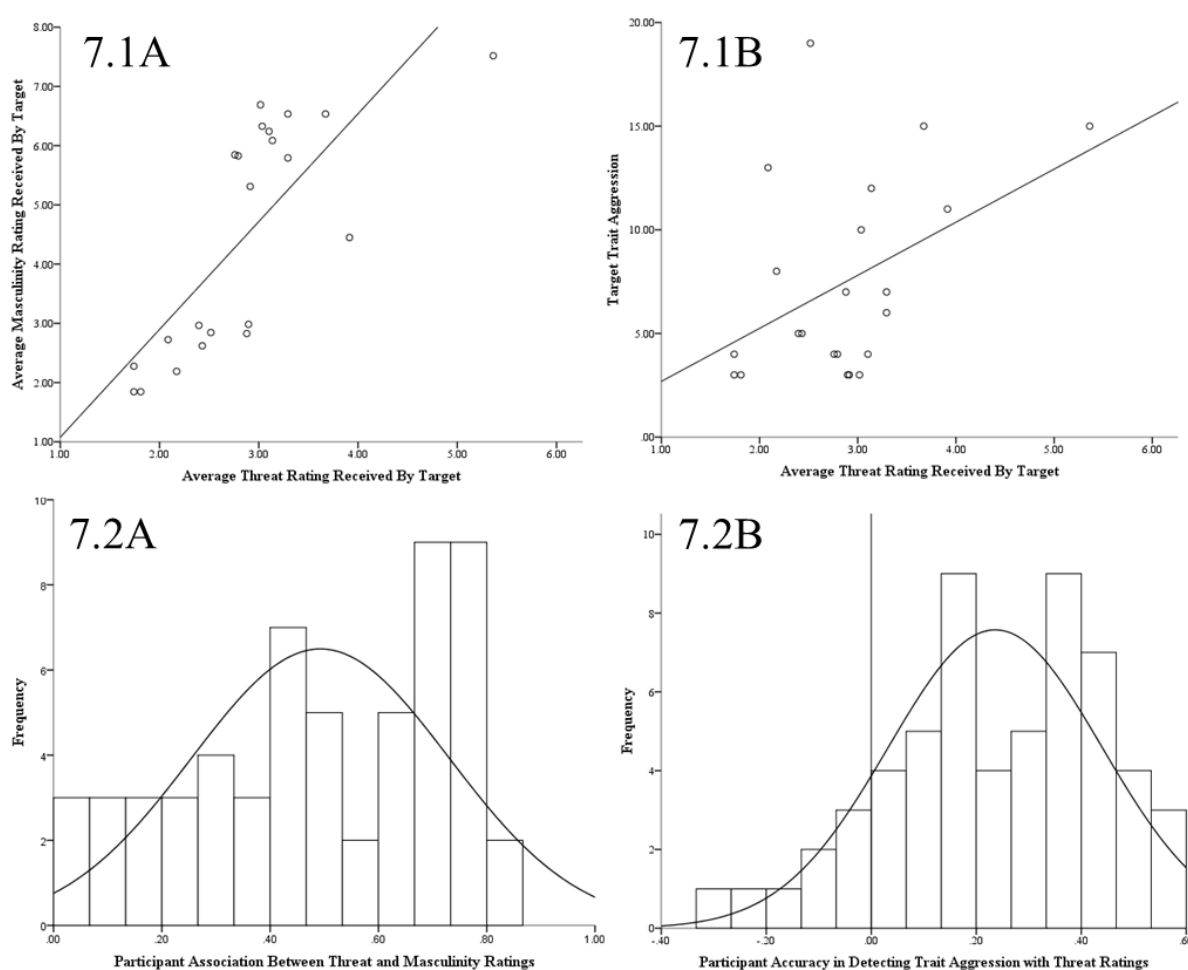


Figure 7.1. The correlation between the average threat rating received by the target and A) the average masculinity ratings received by the target; B) the trait aggression of the target. *Figure 7.2.* The distribution of A) participants' association between threat and masculinity ratings and B) participant accuracy in detecting trait aggression with threat ratings.

7.3.3 Idiographic descriptions of target personality salience; products

standardised threat ratings and aggression. Judge idiosyncrasy is increasingly reported in research (as noted above). However we find no evidence, in experiments of this type, of individual differences in *targets* being reported. In Funder's terms (1995; 1999) accurate judgments of another person's personality are possible if the targets' traits are available to the perceiver. Just as we find individual differences in judges' abilities to detect and/or utilize traits (see Section 7.2); there may be individual differences in the salience of the targets' traits. It could be that one or two targets' demeanours are driving this accuracy.

Traditional Pearson's correlations cannot be used to analyse the recognition of one target's trait from many judges, as multiple judges are making ratings of a single property of the target. Instead, we utilize the core mathematics of the correlation; the average agreement between two sets of values. An efficient measure of agreement is the product of two z scores. In effect, the question is whether a judge rates a target higher on [perception] when that target has a higher [trait relevant to perception]. Here, we want to know whether a judge rates a target as more threatening (than the other targets they observe) when that target *is* more aggressive (than the other targets.) To investigate which target receives a higher threat rating than the other targets or if that target is more aggressive than other targets, judge ratings and target trait scores need to reflect a distribution of the responses. We can demonstrate the distribution of scores by standardising (producing z scores) for the targets' traits (τ) and the judgments made by each judge (j). The standardised scores can efficiently demonstrate if a target has a trait score or threat rating higher than the average target (a $+z$ score) or a lower aggression score or threat rating than the average target (a $-z$ score). The product (multiplication) of these two scores would give an index of agreement between target's aggression and judge rating; $z(\tau)*z(j)$, which we term target Salience (S).

If a target is more aggressive than the average target ($+z(\tau)$) and is judged above average in threat ($+z(j)$), the product of these two numbers would also be positive ($+S$), signifying ready detection of the trait by judges. A target being less aggressive than the average ($-z(\tau)$) and being rated as less threatening than the average target ($-z(j)$) would also demonstrate agreement ($+S$). If there is disagreement between the judgment and trait (if $+z(\tau)$ and $-z(j)$ or $-z(\tau)$ and $+z(j)$) then the product would be negative ($-S$), showing disagreement, demonstrating difficulty in detecting that target's traits. The larger the value of S , the more distinct the target's traits are and the easier participants found it to identify the traits of that target¹³. The average of these S values () can be used to produce information on how well the sample recognised that target's traits. Further, the variation in S values (S_o) per target is indicative of the targets that were more subject to individual variation in judges' abilities; a target that judges cannot easily rate would have a high S_o score.

With information on the individual differences between targets we can test for effects of target properties on salience. We can demonstrate that the sex of the target had little effect on average trait salience of the target ($t(20)=.54, p=.593, d=.37$) and that target trait aggression did not greatly improve average trait salience of the target ($r(22)=.22, 95\% CI [-.54, .59], p=.337$). This analysis also demonstrates that the salience of our targets' trait are highly skewed ($= .23$, $= .62$, $= -.57$, $= 2.58$, Kurtosis= 9.94, Skewness= 2.75, see figure 7.3A). This is in large part to a clear outlier (that target being nearly 4 standard deviations from the mean, $z()= 3.79$). It would seem that participants found it very easy to detect this target's trait aggression. Removing this target 'normalises' the data ($= .11$, $= .34$, $= -.57$, $= .84$, Kurtosis= .36, Skewness= .46). This target was a high scorer on the aggression inventory (scoring 15) but not the highest scorer,

¹³ It is a legitimate concern that S size may also be driven by the size of the target's $z(\tau)$ score, i.e. an outlier in τ may also be an outlier in S . However, a measure of S scaled by τ ($: \tau$) was shown to be strongly congruent with ($r = .73$) and so, for pragmatic reasons, is used for further analysis and advocated for other researchers to use.

so this outlier is not simply a product of being an outlier on the aggression dimension (the other target who scored 15 on aggression received an α value of .84 and the target who scored 19 on aggression received an α value of -.33).

Interestingly the removed target was not one of the targets who elicited large amounts of variation in judges' ability to detect their traits ($\alpha = .70$, $\alpha = .55$, $\alpha = .08$, $\alpha = 2.25$, Kurtosis= 2.34, Skewness= 1.53, see figure 7.3B). A high S_σ score indicates more variation in the match between traits and judgments. This score is informative about the ambiguity of the traits presented by some targets. As ambiguity is to be expected in interpersonal perception research (especially in our paradigm where there are no overt aggressive acts) there is no reason to remove a target who receives a high S_σ score. However, a target with a high α score increases the chance of Type I error in analysis of both the previous analysis sections.

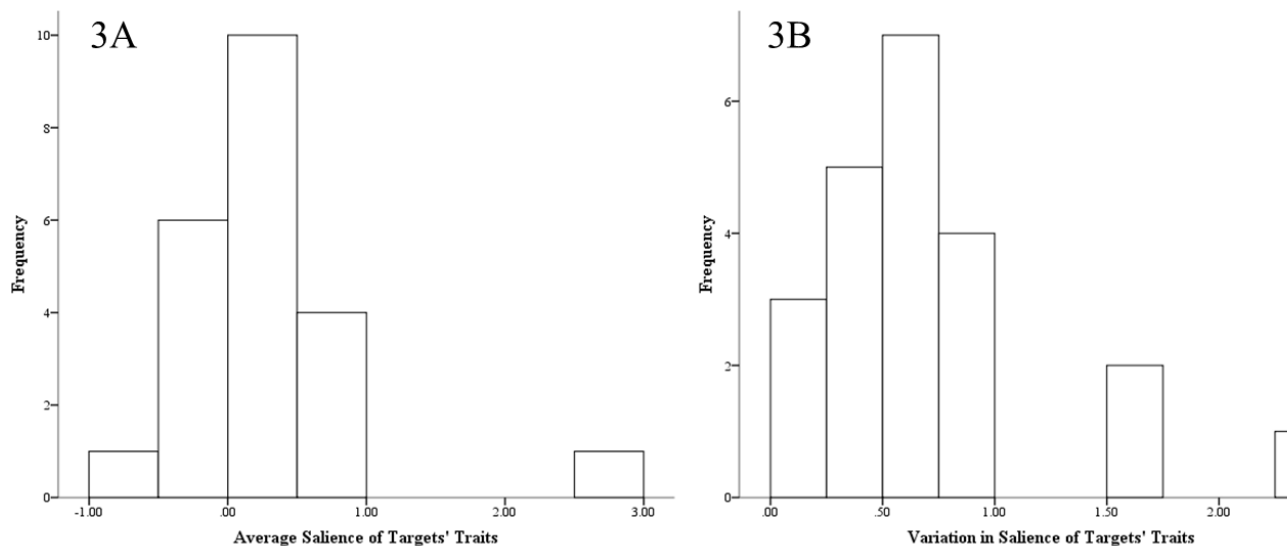


Figure 7.3. The distribution of A) average target salience and B) variance in target salience. Note: the outlier to the right of figure 7.3A and the outlier to the right of figure 7.3B are different targets (see section 3.3).

If we remove the target who was atypically benefitting the judges, we can revisit the original analysis (in section 3.1). When correlating the average threat rating received by the targets with their trait aggression we find that the size of correlation decreases by a surprising amount ($r(21)=.31$, 95% $CI [-.07, .66]$, $p=.169$), although it is still of a meaningful size and the same direction. Revisiting the analysis by judge (in section 3.2) with this salient outlier removed we find that distribution have changed too

($M_{\text{Accuracy(Revised)}} = .13$, $SD_{\text{Accuracy(Revised)}} = .22$, $Min_{\text{Accuracy(Revised)}} = -.38$, $Max_{\text{Accuracy(Revised)}} = .57$, Kurtosis = $-.65$, Skewness = $-.12$), and meaningfully so in a test comparing judge accuracy before and after the outlier has been removed ($t(57)=10.05$, $p<.001$, $d=.50$).

As with all the analyses above, we suggest that this type of analysis adds *descriptive value* to data analysis and all three should be presented when attempting to describe data such as this. Understanding which participants and which targets drive the relationship between all judges and all targets should be of great interest to individual differences psychologists.

7.4. Discussion

Nomothetic analyses of trait judgment studies are sensitive to variations in targets and do not account for variation in judges. Using nomothetic analyses, idiographic judge analyses and idiographic target analyses, the current experiment demonstrates that judges' threat judgments reflect the trait aggression of walking targets. This finding is a replication of previous work (Chapters 3, 4, 5 & 6) and demonstrates that participants are able to make judgments of targets that can reflect (often small) differences in self-report measures made by targets. This accuracy is maintained, even when having to look over one's shoulder at an approaching person.

The main aim of this paper, however, was to *i*) add more weight to the established commentaries of previous authors on idiographic reporting of judges (Brand & Bradley,

2012; Kolar, Funder & Colvin, 1996; Monin & Oppenheimer, 2005) and *ii*) to advocate the idiographic reporting of targets. Nomothetic analysis, as is frequently used in the literature, ignores variation in both judges and targets. Such analyses simultaneously suppress the accuracy of particularly good (or bad) judges (through aggregation of judgments) and bolster the influence of targets who have traits that are atypically salient. It could be the case (as was demonstrated in the results here) that a single target can ‘improve’ the reported accuracy of a nomothetic analysis. It is not possible to identify particularly salient *a priori* as the very nature of Salience is only elicited in the results. Therefore, only *post hoc* idiographic target analysis can reveal any disproportionately Salient targets and allow for better data cleaning. It is our hope that other authors will report more information on the distribution of their targets as this will give a better description of the features driving the results in many interpersonal perception and trait recognition experiments. Nomothetic analyses still have a place in results sections, reporting the consensus of the whole sample but the individual variation within judges and targets can be more interesting and insightful.

Using the measures presented here further research could better explore features of judge accuracy and target salience. For example, further research could explore our incidental (and cautiously treated) finding of sex differences on the accuracy of judges. It would also be of interest to explore our Salience measure in more detail. It would be important to know what features of an outlier target would make their traits particularly salient.

There may well be other ways of analysing target idiosyncrasy which could be developed by other authors, but our intention in this paper was to demonstrate a simple and effective measure which uses common statistical techniques, such as z scores. More than

anything else, we advocate the publication of more commentaries on the inaccuracies of nomothetic analyses to bolster awareness.

7.4.1 Conclusion. Accurate judgments of another person's traits are dependent on both the judge's ability to recognise traits and the salience of the target's traits. Many extant studies test how the average judge responds to all targets, but it should be the objective of individual differences research to explore the performance of individuals. It is demonstrated in this data set is evidence of variability in both judges and targets in being able to accurately judge and elicit trait aggression information. Nomothetic analyses are vulnerable to target salience and ignore judge variation, and so should be reported and interpreted with caution.

References

- Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of personality and social psychology*, 55(3), 387.
- Back, M. D., Schmukle, S. C., & Egloff, B. (2008). How extraverted is honey. bunny77@hotmail.de? Inferring personality from e-mail addresses. *Journal of Research in Personality*, 42(4), 1116-1122.
- Brand, A., & Bradley, M. T. (2012). More voodoo correlations: when average-based measures inflate correlations. *The Journal of general psychology*, 139(4), 260-272.
- Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological Science*, 20(10), 1194-1198.
- Coetsee, V., Chen, J., Perrett, D. I., & Stephen, I. D. (2010). Deciphering faces: Quantifiable visual cues to weight. *Perception*, 39(1), 51.
- Doll, L. M., Hill, A. K., Rotella, M. A., Cárdenas, R. A., Welling, L. L., Wheatley, J. R., & Puts, D. A. (2014). How well do men's faces and voices index mate quality and dominance?. *Human Nature*, 25(2), 200-212.
- Feinberg, D. R., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2008). The role of femininity and averageness of voice pitch in aesthetic judgments of women's voices. *Perception*, 37(4), 615.
- Fong, K., & Mar, R. A. (2015). What Does My Avatar Say About Me? Inferring Personality From Avatars. *Personality and Social Psychology Bulletin*, 41(2), 237-249.
- Fruhen, L. S., Watkins, C. D., & Jones, B. C. (2015). Perceptions of facial dominance, trustworthiness and attractiveness predict managerial pay awards in experimental tasks. *The Leadership Quarterly*.

- Funder, D. C. (1980). On seeing ourselves as others see us: Self-other agreement and discrepancy in personality ratings. *J Personality*, 48(4), 473–493.
doi:10.1111/j.1467-6494.1980.tb02380.x
- Funder, D. C. (1995). On the accuracy of personality judgment: a realistic approach. *Psychological review*, 102(4), 652.
- Funder, D. C. (Ed.). (1999). *Personality judgment: A realistic approach to person perception*. Academic Press.
- Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, 21(3), 177-182.
- Gillath, O., Bahns, A. J., Ge, F., & Crandall, C. S. (2012). Shoes as a source of first impressions. *Journal of Research in Personality*, 46(4), 423-430.
- Hall, J. A., Pennington, N., & Lueders, A. (2013). Impression management and formation on Facebook: A lens model approach. *New Media & Society*, 1461444813495166.
- Harms, P. D., Han, G., & Chen, H. (2012). Recognizing Leadership at a Distance A Study of Leader Effectiveness Across Cultures. *Journal of Leadership & Organizational Studies*, 19(2), 164-172.
- Harrison, M. A., Shortall, J. C., Dispenza, F., & Gallup, G. G. (2011). You must have been a beautiful baby: Ratings of infant facial attractiveness fail to predict ratings of adult attractiveness. *Infant Behavior and Development*, 34(4), 610-616.
- Hirschmüller, S., Egloff, B., Schmukle, S. C., Nestler, S., & Back, M. D. (2015). Accurate judgments of neuroticism at zero acquaintance: A question of relevance. *Journal of personality*, 83(2), 221-228.
- Holleran, S. E., & Mehl, M. R. (2008). Let me read your mind: Personality judgments based on a person's natural stream of thought. *Journal of Research in Personality*, 42(3), 747-754.

- Holtzman, N. S. (2011). Facing a psychopath: Detecting the dark triad from emotionally-neutral faces, using prototypes from the Personality Faceaurus. *Journal of Research in Personality*, 45(6), 648-654.
- Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of personality*, 64, 311-338.
- Lippa, R. A., & Dietz, J. K. (2000). The relation of gender, personality, and intelligence to judges' accuracy in judging strangers' personality from brief video segments. *Journal of Nonverbal Behavior*, 24(1), 25-43.
- Little, A. C., Jones, B. C., DeBruine, L. M., & Dunbar, R. I. (2013). Accuracy in discrimination of self-reported cooperators using static facial information. *Personality and Individual Differences*, 54(4), 507-512.
- McCall, C., & Singer, T. (2015). Facing off with unfair others: Introducing proxemic imaging as an implicit measure of approach and avoidance during social interaction. *PloS one*, 10(2), e0117532.
- McCall, C., Blascovich, J., Young, A., & Persky, S. (2009). Proxemic behaviors as predictors of aggression towards Black (but not White) males in an immersive virtual environment. *Social Influence*, 4(2), 138-154.
- McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology*, 43(2), 293-303. doi:10.1037/0022-3514.43.2.293
- Monin, B., & Oppenheimer, D. M. (2005). Correlated averages vs. averaged correlations: Demonstrating the warm glow heuristic beyond aggregation. *Social Cognition*, 23(3), 257-278.

- Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Personality judgments based on physical appearance. *Personality and Social Psychology Bulletin*.
- Nestler, S., Egloff, B., Küfner, A. C., & Back, M. D. (2012). An integrative lens model approach to bias and accuracy in human inferences: hindsight effects and knowledge updating in personality judgments. *Journal of personality and social psychology*, 103(4), 689.
- Neumann, R., & Strack, F. (2000). Approach and avoidance: the influence of proprioceptive and exteroceptive cues on encoding of affective information. *Journal of personality and social psychology*, 79(1), 39.
- Rule, N. O., & Ambady, N. (2008). The face of success inferences from chief executive officers' appearance predict company profits. *Psychological Science*, 19(2), 109-111.
- Savani, K., Stephens, N. M., & Markus, H. R. (2011). The unanticipated interpersonal and societal consequences of choice victim blaming and reduced support for the public good. *Psychological Science*.
- Semin, G. R., & Smith, E. R. (2013). Socially situated cognition in perspective. *Social Cognition*, 31(2), 125-146.
- Sim, S. Y. L., Saperia, J., Brown, J. A., & Bernieri, F. J. (2015). Judging attractiveness: Biases due to raters' own attractiveness and intelligence. *Cogent Psychology*, 2(1), 996316.

8. General Discussion

8.1 Overview of Main Aims and Findings

This PhD thesis investigates people's accuracy for detecting trait aggression in approaching targets. Specifically, this research adopts and expands the Realistic Accuracy Model (RAM; Funder, 1999) and investigates how features of a walking target (the *relevance* and *availability* of gait information) and the skill of judges (their ability to *detect* and *utilize* gait information) enable accurate trait judgments. The focusing on the RAM, the thesis explored: *i*) the *relevance* of gait information to trait aggression; *ii*) the importance of gait information being *available* to a judge of trait aggression; *iii*) the *detection* of gait information by judges; *iv*) the factors that influence accurate *utilization* of gait information; and *v*) methods of analyses that best describe the RAM components.

The first experiment explored the *relevance* of an individual's style of gait to trait aggression. A biomechanical analysis of thorax and pelvis movement showed that trait aggression can be related to the way an individual walks (Chapter 2). The next two experiments investigated the importance of the *availability* of gait information. In a pair of studies it was shown that *i*) when observing walking targets, observers' judgments of threat accurately relate to the trait aggression of the targets and *ii*) judgments of threat are most accurate when a target's motion is the only information available compared to when only static information is available (Chapter 3). Chapter 4 aimed to investigate judges' *detection*¹⁴ of trait relevant information. An eye-tracking experiment demonstrated that the judges who were more accurate at judging trait aggression were those who spent more time observing gait-relevant aspects of targets (more than targets' faces). Age and life

¹⁴ By the nature of the concepts of detection and utilization, it is difficult to experimentally separate the recognition and the application of stimulus information to a judgment. As Funder notes; "the distinction between detection and [...] utilization, is the same as the distinction between perception and cognition, and it is just as blurry." (1999, p.131). The results presented in Chapter 4 demonstrated that gaze behaviour, which is intrinsic to 'detection', affects judgment accuracy. However, it is important to note that Funder's 'detection' is not simply the mechanical aspects of gaze but the recognition of the observed information.

experience could be important factors in how judges *utilize* gait information to reach a judgment about trait aggression. Chapter 5 demonstrated that accuracy in detecting trait aggression improves after 18 years of age. Chapter 6 showed that accuracy is maintained into older age (60-91 years). The RAM acts as a framework for conceptualising accuracy, as it highlights the importance of both judges and targets in trait judgment research.

Despite this attention to the importance of targets in the RAM, past research has usually used analyses of trait judgments that do not account for target variation (see critiques in Chapter 7; Brand & Bradley, 2012; Kolar, Funder & Colvin, 1996; Monin & Oppenheimer, 2005). Chapter 7 aimed to provide a novel approach to analysing trait judgment data. The results demonstrated that there is notable variation in the ability of judges to accurately judge trait aggression as well as variation in the salience of targets' traits. Chapter 7 highlights that, if the target is to be considered theoretically important, then the target should also be considered analytically important.

In sum, the seven experiments presented in this thesis offered valuable replications of the same finding that, for judges over the age of 18 years, threat and intimidation ratings of moving targets accurately predicted the self-reported trait aggression of those targets. Moreover, the most important information for the accurate detection of threat were those associated with the gait of the targets. This programme of research is also the first attempt to systematically investigate the different aspects of the RAM and demonstrates clear individual differences (in targets and judges) in accurate judgments of aggression.

8.2 Theoretical Implications

The thesis draws on a broad background of research areas, touching on theories of approach and avoidance behaviour, aggression, trait recognition and interpersonal affordances. The thesis as a whole makes significant contributions to both the personality and aggression detection literature and these are outlined below.

8.2.1 Implications for personality psychology. Chapter 2 refers to the philosopher Charles Taylor who states; "...the most pervasive features of my attitude to the world and to others is encoded in the way I project myself in public space" (1995, p.171), and this thesis provides clear evidence to support his argument. It is a novel finding that trait aggression is salient to judges simply from watching a person walk (Chapters 3, 4, 5, 6 & 7) and that individuals' gait biomechanics relate to their trait aggression (Chapter 2). In short, the current work serves as the first evidence of a relationship between personality, gait biomechanics and judgments of an individual in motion.

Previously it was shown that personality judgments from gait were *inaccurate* and that gait biomechanics did *not* relate to personality attributes (Thoresen, Vuong & Atkinson, 2012). The difference between the findings presented here and those of Thoresen and colleagues could be the result of the current focus on aggression rather than Thoresen's focus on the Big Five personality traits. The difference in results could also be the product of differing biomechanical measures (see Chapter 2) and the way in which judgments were collected.

The participants who took part in the perception experiments for this thesis were asked to perform a different task from those who were recruited for Thoresen et al's research. In the current work, participants were asked how threatening or intimidating they found the targets to be. In effect they were indicating how the target stimuli made them feel (threatened/intimidated). The participants in the Thoresen et al. study were asked to rate targets on traits such as extraversion and neuroticism. They were not considering how the targets made them feel. It is possible that accuracy improves when judgments directly impact upon the judge. It may be simpler for participants to respond with an accurate judgment about how a target makes them feel (e.g. threatened) rather than judging the traits of others (e.g. extraversion).

The experiments presented in this thesis mirror the components of Funder's (1999) Realistic Accuracy Model (RAM). Funder proposes that accurate judgments of another's personality are based on that target person exhibiting behaviours that are relevant to their traits. When these *relevant* behaviours are *available* to a perceiver, that perceiver can *detect* those behaviours and *utilize* the information to make an accurate judgment. As many previous studies do not explore the availability, detection or (factors affecting) utilization of target information, the programme of research presented here is the first to actively explore all aspects of the RAM on a single topic. This thesis demonstrates that an individual's gait is *relevant* to their trait aggression. When motion information is not *available*, accuracy at detecting aggression is poorer and, in fact, there are targets that appear to have more available traits than others (see Chapter 7). For the current thesis, those participants who spent more time observing gait-relevant aspects of targets were more accurate at detecting trait aggression. This finding, it could be argued, is evidence that those participants were prioritising the *detection* of gait when making their threat judgments. *Utilization* is harder to investigate however Chapter 5 demonstrated that adult participants were more accurate at detecting threat than child participants. This finding suggested that, with age, comes the ability to better utilise the information presented to a perceiver. Further, the current findings demonstrated individual differences in the *detection* and/or *utilization* of the same information in the form of idiosyncratic judge analysis and variability in ratings received by targets. In sum, the work presented in this thesis is supportive of the RAM as a framework for interpersonal trait recognition. However, as stated in the General Introduction, it is my contention that the RAM could include aspects of Gibson's (1979) theory of affordances. In fact the socialisation of Gibson's affordances (Costall, 1995) offers an explanation for the improvement of aggression detection accuracy

with age. Engaging in a more risky, adult culture with peers who can successfully detect relevant traits may increase peoples' ability to detect traits relevant to aggression.

Further research is needed to effectively demonstrate this proposition, but the work in this thesis contributes to the understanding of interpersonal trait recognition as a product of the detection and utilization of interpersonal affordances.

8.2.2 Implications for aggression detection research. As is repeated throughout the thesis, research into judgments of the malevolent traits of others typically uses photographs of faces as representations of other people (e.g. Boshyan, Zebrowitz, Franklin, McCormick and Carré, 2013; Carré, McCormick & Mondloch, 2009; Geniole, Molnar, Carré & McCormick, 2014; Hehman, Flake & Freeman, 2015; Hehman, Leitner & Gaertner, 2013). Some studies present targets briefly (39ms; Carré et al., 2009), others manipulate the features of the face to increase or decrease malevolent perceptions (Hehman et al., 2013) and others test for accuracy of aggression judgments (such as; Boshyan et al., 2013; Carré et al, 2009; Geiniole et al., 2014). In all of these cases, the findings of the studies are interesting in that, from brief presentations of photographs, participants can make reliable judgments of another person's posed intimidation or aggression. Furthermore, research has demonstrated that these perceptions can be accurate. However, these studies do not use realistic stimuli. It is rare to engage with a person who appears, for microseconds, simply as a face. Whilst it is impressive that reliable and/or accurate judgments of others can be made at such speed, caution must be exercised when applying such findings to everyday life. Some threat judgment studies have used whole body photographs of stimuli (see; Azarian, Esser & Peterson 2015a; Azarian, Esser & Peterson, 2015b). However this past research still used static presentations of individuals in unusual poses (e.g Azarian et al. used photographs of full body enactments of emotions which were arguably overacted).

What is lacking from the extant literature, and importantly, what this programme of research contributes to the literature, is realistic presentation of potential dangers. In everyday life, it is not often the case that a potential threat appears for only microseconds nor that an individual expresses an overt, whole body, demonstration of anger. The critical contribution of this thesis is that the (relatively mundane and automatic) act of walking is indicative of the potential danger another person poses. An approaching person is likely to be visible for a significant period of time and that person is *moving*. It may well be the case that movement, more so than a still image, is informative when making judgments about the danger another poses in both lab (see the direct comparison to Boshyan et al.'s, 2013, work in Chapter 6) and real world settings. This thesis demonstrates that gait is important in detecting trait aggression and that whole body movement should be considered more often in research where the perceived malevolence of others is under investigation.

8.3 Practical Applications

The practical applications of the findings of this thesis generally form two categories; proximal and distal detection of dangers. Individuals may benefit from information and training on how to detect proximate dangers, when making interpersonal judgments of aggression and professional security services may benefit from understanding how to detect threats from a distance via CCTV or security networks.

8.3.1 Benefits for interpersonal judgments of aggression. Asynchrony, between interpersonal judgments of potential dangers and the potential danger genuinely posed by another person, can be costly. When individuals are over-cautious, assuming people are threats when they are not, they may restrict their behaviours and actions, such as can be seen in surveys of adults' deterrents to being active (Dawson, Hillsdon, Boller & Foster, 2007; Foster, Knuiman, Hooper, Christian & Giles-Corti, 2014a; Kramer, Maas, Wingen & Kunst, 2013), and in parents' restrictions of children's independent behaviours (Foster,

Villanueva, Wood, Christian, & Giles-Corti, 2014b). When individuals are under-cautious about the danger another person poses they could become a victim of crime. However, the evidence presented in this thesis suggests that individuals over the age of 18 years old can accurately detect individuals who are aggressive. It is, therefore, likely that problematic issues with the detection of interpersonal threat are the product of metacognitive self-doubt. That is, an individual may either recognise an aggressive individual but dismiss this belief or not recognise their recognition. Both of these potential explanations of inaccurate judgments are based on the context of the situation. In the standard, safe, experimental setting used routinely throughout this thesis, participants had neither time pressure nor environmental influences to inflate or deflate their feelings of threat. This may have facilitated accurate judgments of threat. On a street late at night, in situations where anxiety may be heightened, accurate threat perceptions may be less likely. Regardless of the potential environmental effects, this thesis demonstrated that individuals *can* be accurate in their judgments of others, and therefore it is possible that this accuracy could be recognised, trained and developed in individuals who may be over or under cautious of others. Such training could use the evidence presented here, for example the visual search style of accurate detectors of aggression (Chapter 4) and the biomechanics of an aggressive gait (see Chapter 2), to enhance the accuracy of judges. Through the demonstration, via a training course, of an individual's accuracy in detecting potential aggressors, it could be possible to build their recognition of, and belief in, their accuracy. This could be done simply using the experimental paradigm presented throughout this thesis with trainees judging video recordings of target stimuli and receiving feedback regarding their threat detection accuracy. Increasingly, simulated, 'safe' environments are being recommended for training skills such as performance in ball sports (Miles, Pop, Watt, Lawrence, & John, 2012) and surgery (John, 2008) and it is reasonable to expect that similar 'practice

environments' could benefit confidence (and perhaps even accuracy) in detecting aggressive others. In fact skill efficiency greatly improves with practice and feedback (borrowing, again, from the medical practice literature; Bosse et al., 2015) so it could be possible to coach those who feel intimidated by strangers to recognise their ability to discern genuine aggressors.

8.3.2 Benefits for detecting aggressors from afar. Beyond the interpersonal level, there is good reason to suggest that the research presented in this thesis would benefit professionals whose job it is to detect impending crimes. Troscianko et al. (2004) conducted an experiment where professional CCTV observers and naïve participants were asked to detect the likelihood of an observed event leading to a crime. Their critical findings are all of interest to the conclusions presented in this thesis. First, their participants could detect impending crimes. Second, there was no real difference in performance between professional and naïve observers' accuracy, all participants performed well. Third, behaviours such as gait and posture appeared instrumental in predicting criminal activity. It is important to note that Troscianko et al. used many different crime types, not simply interpersonal aggressive crimes. Their finding that observers were able to detect impending criminal acts is similar to the finding presented throughout this thesis; that adults are able to detect dispositional aggression in others from observations of walking behaviour. Troscianko and colleagues claimed that the accurate predictions made by their participants were facilitated by the behaviours of the targets in the CCTV videos, including the targets having a 'distinctive gait'. Although they did not elaborate on what a distinctive gait was, it would be no surprise if this gait was one that was high in relative thorax to pelvis range of movement (given the evidence presented in Chapter 2). The research presented in this thesis extends Troscianko et al.'s work; demonstrating how certain movements can elicit information relevant to dispositional

aggression and the specific ways that individuals search for these types of movements. As with the ‘street level’ interpersonal aggression training suggested above, the results of the studies throughout this thesis could be used to train those with a professional interest in detecting individuals who pose a high risk of aggression. These include CCTV operators, the police, border control personnel, security professionals and door staff or ‘bouncers’ who assess risks to ‘night time leisure venues’ (see Hobbs, Hadfield, Lister & Winlow, 2002). Whilst there may be more overt and relevant cues to aggressive individuals in such situations (such as evidence of alcohol consumption, Hobbs et al., 2002, or affiliation with ‘football firms’, Ayres & Treadwell, 2011) there may be situations where detection of potential aggressors is benefitted by using behavioural cues such as gait.

8.4 Methodological Limitations

The majority of the experiments in this thesis used a very similar method of collecting data; participants observed presentations of target people and made ratings of the targets on rating scales. As such, these experiments had similar methodological limitations. Generally these limitations can be classified into; *i*) the limitations of the method for demonstrating the targeted theoretical implications and *ii*) the way ‘ground truth’ of aggression of the targets was established.

8.4.1 New stimuli, same old paradigm. This research programme extended previous research into judgments of the others’ aggression by focusing on judgments of targets made from observing full body movements. It is written throughout the thesis that presenting a dynamic whole body stimulus is an improvement on the typical paradigm where participants make judgments of targets presented as photographs of faces. Such experiments over-represent the magnitude of influence the face has on judgments of malevolent attributes by limiting the information available to the perceiver (such as, Boshyan, et al., 2013; Carré, et al., 2009; Geniole, et al., 2014; Hehman, et al., 2015;

Hehman, et al., 2013). However, it is important to note that by standardising the clothing worn by the targets, standardising a laboratory environment background for the videos and encouraging targets to present themselves with a neutral facial expression, the studies in this thesis were, arguably, also guilty of restricting available information. It is clear that presenting the target as a video of the whole body walking is more informative (and perhaps more ‘ecologically valid’) than presenting a target as a photograph of their face, but research has demonstrated that a target’s facial expression (Taylor & Barton, 2015) and clothing (for a review, see; Elliot & Maier, 2014) can influence perceived malevolence. It is perhaps the case that accuracy in judging potential dangers may increase (due to clothing reflecting the wearer’s disposition) or decrease (where an individual’s dress style appears different to their disposition) with variation in clothing. Facial expressions and displays of emotion in general could allow better discrimination of impending acts of aggression, when an individual is about to act in anger (see; Azarian, et al. 2015a; 2015b). It is worth noting, however, that part of the novelty of the current work is that judgments of others can reflect the disposition of others from a relatively neutral movement, with long term trait information being communicated without immediate, situational, cues to aggression.

This thesis tentatively makes reference to the potential of behavioural *affordances* as an explanation for how interpersonal accuracy is possible. The suggestion is made cautiously due to the lack of active perceivers in the experiments presented here. Funder (1999) claims that Gibson’s model of affordances does not apply to detecting the traits of others, due to the fact that many accurate participants in laboratory experiments are passive perceivers (not Gibson’s ‘active perceivers’). The current research also forced participants to be ‘passive’, simply observing videos presented on a screen. It is plausible that an individual may act and react towards an approaching person, changing gait style in response to the style of an approaching person or attempting to ‘probe’ interpersonal

information from a stranger with smiles or bowed heads. It is a limitation of the current work that participants had no chance to engage reciprocally with targets. This particular limitation is not uncommon and Ulric Neisser noted this same issue in 1980; “the theories and experiments described here all refer to an essentially passive onlooker, who sees someone do something (or sees two people do something and then makes a judgment about it. He (this is the generic passive *he*) doesn’t *do* anything ---- doesn’t mix it up with the folks he’s watching, never tests his judgments in action of interaction” (Neisser, 1980, p603, emphasis and punctuation from original). If participants had had the opportunity to attempt behavioural ‘probes’ in this experiment, if they had a chance to act and react towards the targets, they may have made even more accurate judgments of the targets’ trait aggression.

Neisser’s (1980) concerns about the state of social psychology methods go beyond the absence of interaction. He also shows concern about the nature of judgments based on Likert rating scales. Putting a number to a social judgment (‘this person is 6 threatening’) is a highly artificial paradigm that has become the cornerstone of social judgment research. There is a literature that raises concerns with how participants use Likert scales in general. For example, individual participants in a sample might consider the poles of the Likert scale in different ways to each other. For example participants might see responses on a 1-9 Likert scale of threat meaning different things. They could see 1 as the least threatening person they have ever seen, the least threatening person in the current study or in a potential given interaction context (that the participant imagines) that person would never pose a threat. The potential influences caused by using Likerts comparatively (both within and/or outside the current study) and relative to oneself, are known as ‘anchoring’ effects (Mussweiler, & Strack, 1999; Tversky, & Kahneman, 1974; Wilson, Houston, Etling, & Brekke, 1996). Anchoring effects can have a large influence on nomothetic data analysis

(as used in Chapter 3). If, in a given sample, a large minority of judges anchor all their threat judgments between 1 and 3 (from a possible scale of 1 to 9) then the aggregate judgment of ‘threat’ given to a target will be disproportionately lower. However, it is a strength of idiographic judge analyses that they are more resistant to these effects (See Chapters 4, 5, 6 & 7). The idiographic judge analysis of those judges who are anchoring their judgments between 1 and 3 will still reflect the accuracy of those anchored responses.

Throughout this thesis, the recognition of aggression through threat ratings is considered a ‘trait’ of judges. Judges are frequently referred to as being more or less accurate in their judgments in general (in a manner typical of the aggression detection literature; Boshyan et al., 2013; Carré et al, 2009; Geiniolo et al., 2014). However, for an ability to be considered a trait of a person it should be demonstrated to be *stable* over time. There is a notable absence of test-retest studies on social judgments using Likert scales. This may be due to many studies not being interested in the stability of a judgment, for example a first impression judgment needs only be accurate in that moment, but this thesis is implicitly treating judgment accuracy as an ability of the judges. Future research could easily address this gap in the literature with a test-retest judgment paradigm. The lack of test-retest measures in this thesis is a consequence of using new presentation formats of stimuli but the same judgment old paradigm.

8.4.2 Quantifying aggression. It is only possible to measure ‘accuracy’ in detecting trait aggression with a good ‘ground truth’ measure of targets’ aggression. In all the studies included in this thesis, targets’ aggression was measured using the Buss-Perry Aggression Questionnaire (Buss & Perry, 1992). The questionnaire has been shown to be a valid measure of current aggression (Bryant & Smith, 2001, see Appendix F), aggression in a hypothetical context (Archer & Webb, 2006; O’Connor, Archer & Wu, 2001) and historic aggression (Diamond, 2006) and has been used with both student (García-León et

al., 2002) and forensic (Diamond, Wang & Buffington-Vollum, 2005) populations. However, the scale remains a self-report measure of aggression, dependent on both participants' honesty in reporting and their ability to introspect their own dispositional aggression. It is, in a way, a moot point to criticise the validity of the measurement of aggression in this context, as judges were routinely able to accurately detect the trait aggression of the targets. If participants had made no accurate judgments then it would be evidence of no relationship between movement and aggression. However, as participants were accurate they were recognising an attribute of the targets from the targets' gait which was also reflected in the self-report measure.

Perhaps the best reason for continuing to use a self-report measure of trait aggression is the lack of an alternate, ethical, and experimental measure of aggression. Tedeschi and Quigley (1996) questioned the construct validity of experimental measures of aggression, stating that the most common forms of laboratory-measured aggression typically misrepresent the motive of the aggressive act, use retaliatory paradigms of aggression more than instigative acts and often use measurements which record socially acceptable forms of aggression (such as 'rough and tumble play') more than true aggression. They cite methods such as the 'Teacher/Learner' paradigm (where participants are "given a chance to retaliate against the confederate, usually by delivering shocks to him or her"; Tedeschi & Quigley, 1996, p. 165) and the 'Bobo Modelling' paradigm (which measures child imitation of adults' aggression towards an inflatable toy) as examples of common but not valid measures of aggression. Ritter and Eslea (2005) have also critiqued the validity of more recent measures of aggression and suggest that Tedeschi and Quigley's (1996) arguments still stand even with more recent, experimentally derived, measures of aggression. The more modern methods include the 'Hot Sauce' paradigm ("participants are required to determine the amount of hot sauce to be (purportedly)

consumed by another person who allegedly does not like spicy foods”; Ritter & Eslea, 2005, p. 410) or the ‘Bungled Procedure’ paradigm (“participants are given the opportunity to shoot at a human target with a pellet or paintball gun.[...] Aggression in this paradigm is operationalised as the power of the gun chosen [...] multiplied by the number of pellets elected to be used to shoot at the target”; Ritter & Else, 2005, p. 411). Giancola and Chermack (1998) criticised Tedeschi and Quigley’s (1996) paper, and promoted the use of the Point Subtraction Aggression Paradigm (PSAP) for laboratory measures of aggression. The PSAP (see; Golomb, Cortez-Perez, Jaworski, Mednick & Dimsdale, 2007) is well used in the aggression literature (e.g. Boshyan, et al., 2013; Carré, et al., 2009; Geniole, et al., 2014) as a measure of ground truth in various investigations into judgments of aggression from presentations of faces.

Golomb et al. (2007) described the PSAP as; “measuring the number of "aggressive" responses toward a fictitious partner in a computer-driven scenario. Subjects press one button to accrue points that will translate to earned money and press a different button to subtract money from a person with whom they are partnered, at no gain to themselves. These occur in the context of scheduled subtraction of earned points - and resulting subtraction of monetary earnings - that are attributed to the fictitious partner. Point subtractions by the subject are considered to represent aggressive responses.” (p. 95-96). The PSAP has been shown to relate to behavioural violence (Cherek, Moeller, Schnapp & Dougherty, 1997) and is frequently used in neurological and endocrine research into aggression (Bubenzer-Busch et al., 2015; Cote, McCormick, Geniole, Renn & MacAulay, 2013). However, this paradigm is highly artificial and, in a real world context, would appear to be more indicative of tendencies to acquisitive (‘stealing’) offences than aggression. Further, the PSAP fulfils all three of Tedeschi and Quigley’s (1996) criticisms; *i*) the motive is atypical (the paradigm is usually driven by financial

gains), *ii*) the PSAP is a measure of retaliatory aggression and *iii*) the PSAP is a socially sanctioned aggressive act. Even this paradigm, one of the most popular methods of experimentally measuring aggression, has not provided sufficient validity to be considered an improvement on self-report tools.

All of the experimental paradigms above struggle to tap into spontaneous, interpersonal aggression. By the very nature of the construct being measured, it would be nigh-impossible to ethically and experimentally capture interpersonal violence in a laboratory setting. Whilst relying on a self-report measure of trait aggression is limiting, questionnaire methods allow participants to tap into experiential and introspective information about their inclination to aggression. In fact, given the lack of alternatives that adequately imitate real world aggression, it is reasonable to advocate the use of self-report measures of dispositional aggression. Further, as much of the literature and theory for this thesis is sourced from personality psychology, it is not atypical to consider the ‘ground truth’ of an individual’s traits as those they self-report on a personality questionnaire (Funder, 2012; Funder & West, 1993). In future research, it would be of interest to recruit targets who genuinely pose a danger to others perhaps by filming the gait of violent offenders. However, given the limited accessibility of such populations and genuine acts of aggression, researchers should have good reason to be content with using self-report aggression measures.

8.5 Future Research

The research presented in this thesis has acted as a ‘proof of concept’ for a new direction in aggression detection research, incorporating concepts from ecological, developmental and personality psychology. Each experiment has prompted ideas for future work. Future research should endeavour to expand on the cross section of ideas presented in the thesis; *i*) developing more realistic contexts for threat judgments, *ii*) explaining the

influence of specific aspects of targets on judgments and *iii*) exploring the nature of polyadic judgments of groups of targets.

8.5.1 Exploring the context of judgments. As with most laboratory research, the paradigm used throughout this thesis is somewhat artificial. It is unlikely that an individual would be asked to judge the risk (on a rating scale) posed by a walking person presented on a screen for 5 to 10 seconds. Future work should do more to make the context of threat judgments more typical of a risk rich (real world) environment. For instance, manipulations of target clothing could be manipulated to assess their impact on the accuracy of threat. It would be advisable to manipulate (*does a set of clothes affect threat more than another set of clothes?*) and naturalistically vary (*are the clothes a target chooses to wear relevant to their trait aggression?*) the clothing worn by the targets.

Furthermore, the environment in which the danger judgment takes place should be considered. As technology becomes more accessible, research into the judgments of the danger posed by another person in a street setting (which this thesis sought to emulate experimentally) could be conducted in a much more realistic manner. In future research participants could walk down real streets (perhaps even wearing eye tracker devices that can now be the size of sunglasses) and make ratings of targets walking towards them (using applications on mobile devices and tablets.) Furthermore, using such methods, it would be possible to investigate the influence of lighting on judgments of threat. It is difficult to film the approach of an individual with the realistic level of darkness that one might experience when encountering a stranger late at night. Night-time may well increase perceptions of threat (or at the very least increase the anxiety of an individual making such a judgment) and is, perhaps, when judgments of danger from others are most often made. If point light presentations of individuals are comparable to the silhouette of an approaching

person, then threat judgments could be surprisingly accurate under such poor lighting conditions

Alcohol could also influence judgments of potential danger. It has been demonstrated that many victims of stranger-homicide had been drinking prior to the attack (37.10%, in Los Angeles; Goodman et al., 1986). Other research shows that significantly more victims of stranger sexual assault had been drinking alcohol than not drinking prior to the assault (Ullman & Brecklin, 2000). Whilst it would be incorrect to suggest that a victim's consumption of alcohol is a factor in them becoming a victim of homicide or sexual assault, it is of importance to explore what effect (if any) alcohol consumption has on individuals' ability to discern whether or not a stranger poses a danger to them. Alcohol has been shown to have an inhibitive effect on recognition of 'threatening' facial expressions (Sripada, Angstadt, McNamara, King & Phan, 2011) and so could also influence the detection of trait aggression in strangers.

8.5.2 Exploring target characteristics. Unfortunately, for the current thesis, it was not possible to produce films of the target stimuli that allowed for optimum analysis of gait biomechanics and optimum quality of presentation to judges. Future research, should investigate, using motion capture clothing, how the biomechanical aspects of a target's gait directly affect the accuracy of threat perceptions made by observers.

Statistical approaches to analysing how qualities of a target may communicate properties of that target do exist and are frequently used. Karelaia and Hogarth (2008) reported that Egon Brunswick's concept of 'lens modelling' (where the qualities of a target are mathematically associated with both the visually salient aspects of a target and the judgments made by a perceiver) has become increasingly popular in human judgment studies. Such analyses would greatly benefit future research into the communicability of gait in trait detection research. For example, Back, Schmukle and Egloff (2010)

demonstrated that ‘speed and energy of body movement’ is an important communicator of an individuals’ extraversion. More than this, they showed that speed of movement positively correlates with judgments of popularity, which explains how judgments of popularity reflect target’s extraversion. Back et al., (2010) also demonstrated that certain qualities of a target (such as ‘fashionable face and hairstyle’) lead to judgments of popularity but do not reflect ratings of extraversion. Using a similar, comprehensive, analysis of the features of an approaching person (including clothing worn, facial expression, height, weight and multiple gait biomechanics features) could help discern exactly which features communicate trait aggression and which features are misleading. Such analysis could lead to a better understanding of the relative contribution of gait (see Chapter 3) and the commonly researched face shape (see; Carré and McCormick, 2008) to detecting dispositional aggression in others. Such an analysis would be complimentary to Funder’s (1999) Realistic Accuracy Model, allowing precise sourcing of how *relevant* and *available* information is *detected* and *utilised* by a judge.

8.5.3 Exploring collaborative judgments of threats. There is an absence of research into group judgments of groups of targets. In fact, there is a dearth of interpersonal perception research which involves presenting more than one judge or more than one target simultaneously. Future research should address this imbalance. Such research would replicate real-world situations when acts of aggression (or at the very least, intimidation) are perpetrated in groups (such as football hooliganism and gang crime). It is highly likely that presenting multiple approaching targets at once would lead to a decrease in threat detection accuracy. This could be due to the effort required to detect the relevant aggression information from all targets, or, perhaps, due to gestalt effects of considering the targets as a group rather than as individuals.

Future research could investigate judgments of groups by presenting varying numbers of targets at once and asking participants to state how threatened and/or intimidated they feel. Using this ‘first person’ judgment (*how intimidated are you by this group?*) would be beneficial here, as trying to predict the dispositional aggression on a target-by-target basis (*how aggressive do you think Target A is?*) would be more complex and artificial than the perceiver’s experience of the group as a whole. The accuracy of a judgment like this would be difficult to measure, as group members would affect each other’s propensity to aggression, giving no true, single aggression measure. However, it would be of interest to investigate, perhaps using eye tracking methodology, which targets within a group are most influential when it comes to threat judgments.

It would also be beneficial to investigate collaboration between judges when attempting to discern the traits of others. As mentioned in Chapter 4, future research should study how children engage with their parents to acquire threat detection accuracy. Also of interest would be investigations into how groups of judges reach a polyadic decision on the disposition of an approaching individual (or group). The combination of multiple individuals with individual accuracies (and confidences) in detecting trait aggression could lead to a collective accurate or inaccurate decision.

Whilst such experiments seem academic, they are important in terms of how individuals (or rather groups of individuals) engage with the world on a day to day basis. If judgments of others greatly change when there are two targets, or if two judges reach a decision as a pair that is different from their individual judgments, then the findings of the current work might have a constrained application. On the other hand, future research with groups might demonstrate that judgments of a target group made by a group of judges are comparable to judgments of a lone target made by a lone judge. Either way, such research would aid a better understanding of the applicability of the current work.

8.6 Thesis Conclusion

Individuals can accurately recognise the trait aggression of another person by simply rating how intimidating or threatening they find that person to be. This accuracy in recognising the trait aggression of an approaching person improves after the age of 18 years and is maintained into older age. Threat detection is particularly accurate when individuals are presented with only gait information or when individuals fixate on gait-relevant aspects of an approaching person. In sum, how an individual walks communicates information about their trait aggression to others.

References

- Archer, J., & Webb, I. A. (2006). The relation between scores on the Buss–Perry Aggression Questionnaire and aggressive acts, impulsiveness, competitiveness, dominance, and sexual jealousy. *Aggressive Behavior*, 32, 464–473.
doi:10.1002/ab.20146
- Ayres, T. C., & Treadwell, J. (2011). Bars, drugs and football thugs: Alcohol, cocaine use and violence in the night time economy among English football firms. *Criminology and Criminal Justice*, 12(1), 83–100. doi:10.1177/1748895811422949
- Azarian, B., Esser, E. G., & Peterson, M. S. (2015a). Watch out! Directional threat-related postures cue attention and the eyes. *Cognition and Emotion*, 1–9.
doi:10.1080/02699931.2015.1013089
- Azarian, B., Esser, E. G., & Peterson, M. S. (2015b). Evidence from the eyes: Threatening postures hold attention. *Psychon Bull Rev.* doi:10.3758/s13423-015-0942-0
- Back, M. D., Schmukle, S. C., & Egloff, B. (2010). A closer look at first sight: Social relations lens model analysis of personality and interpersonal attraction at zero acquaintance. *European Journal of Personality*, 25(3), 225–238. doi:10.1002/per.790
- Boshyan, J., Zebrowitz, L.A., Franklin, R.G., McCormick, C.M., & Carré, J.M. (2013). Age similarities in recognizing threat from faces and diagnostic cues. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 69(5), 710–718, doi:10.1093/geronb/gbt054.
- Bosse, H., Mohr, J., Buss, B., Krautter, M., Weyrich, P., Herzog, W., ... Nikendei, C. (2015). The benefit of repetitive skills training and frequency of expert feedback in the early acquisition of procedural skills. *BMC Medical Education*, 15(1), 22.
doi:10.1186/s12909-015-0286-5

- Brand, A., & Bradley, M. T. (2012). More voodoo correlations: when average-based measures inflate correlations. *The Journal of general psychology*, 139(4), 260-272.
- Bryant, F. B., & Smith, B. D. (2001). Refining the architecture of aggression: a measurement model for the Buss–Perry Aggression Questionnaire. *Journal of Research in Personality*, 35(2), 138-167.
- Bubenzer-Busch, S., Herpertz-Dahlmann, B., Kuzmanovic, B., Gaber, T. J., Helmbold, K., Ullisch, M. G., ... Zepf, F. D. (2015). Neural correlates of reactive aggression in children with attention-deficit/hyperactivity disorder and comorbid disruptive behaviour disorders. *Acta Psychiatrica Scandinavica*, n/a–n/a.
doi:10.1111/acps.12475
- Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of personality and social psychology*, 63(3), 452.
- Carré, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behavior in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 275, 2651–2656. doi:10.1098/rspb.2008.0873
- Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial Structure Is a Reliable Cue of Aggressive Behavior. *Psychological Science*, 20(10), 1194–1198.
doi:10.1111/j.1467-9280.2009.02423.x
- Cherek, D. R., Moeller, F. G., Schnapp, W., & Dougherty, D. M. (1997). Studies of violent and nonviolent male parolees: I. Laboratory and psychometric measurements of aggression. *Biological Psychiatry*, 41(5), 514–522. doi:10.1016/s0006-3223(96)00059-5
- Costall, A. (1995). Socializing affordances. *Theory & Psychology*, 5(4), 467-481.

- Cote, K. A., McCormick, C. M., Geniole, S. N., Renn, R. P., & MacAulay, S. D. (2013). Sleep deprivation lowers reactive aggression and testosterone in men. *Biological Psychology*, 92(2), 249–256. doi:10.1016/j.biopsycho.2012.09.011
- Dawson, J., Hillsdon, M., Boller, I., & Foster, C. (2007). Perceived barriers to walking in the neighbourhood environment and change in physical activity levels over 12 months. *British Journal of Sports Medicine*, 41(9), 562–568. doi:10.1136/bjism.2006.033340
- Diamond, P. M. (2006). The Short-Form Buss-Perry Aggression Questionnaire (BPAQ-SF): A Validation Study With Federal Offenders. *Assessment*, 13, 227–240. doi:10.1177/1073191106287666
- Diamond, P. M., Wang, E. W., & Buffington-Vollum, J. (2005). Factor structure of the Buss-Perry Aggression Questionnaire (BPAQ) with mentally ill male prisoners. *Criminal Justice and Behavior*, 32, 546-564. doi:10.1177/0093854805278416
- Elliot, A. J., & Maier, M. A. (2014). Color Psychology: Effects of Perceiving Color on Psychological Functioning in Humans. *Annu. Rev. Psychol.*, 65(1), 95–120. doi:10.1146/annurev-psych-010213-115035
- Foster, S., Knuiman, M., Hooper, P., Christian, H., & Giles-Corti, B. (2014a). Do changes in residents' fear of crime impact their walking? Longitudinal results from reside. *Preventive medicine*, 62, 161-166. doi:10.1016/j.ypmed.2014.02.011
- Foster, S., Villanueva, K., Wood, L., Christian, H., & Giles-Corti, B. (2014b). The impact of parents' fear of strangers and perceptions of informal social control on children's independent mobility. *Health & place*, 26, 60-68. doi:10.1016/j.healthplace.2013.11.006

- Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, 21(3), 177-182. doi: 10.1177/0963721412445309
- Funder, D. C. (Ed.). (1999). *Personality judgment: A realistic approach to person perception*. Academic Press.
- Funder, D. C., & West, S. G. (1993). Consensus, Self-Other Agreement, and Accuracy in Personality Judgment: An Introduction. *Journal of Personality*, 61(4), 457–476. doi:10.1111/j.1467-6494.1993.tb00778.x
- García-León, A., Reyes, G. A., Vila, J., Pérez, N., Robles, H., & Ramos, M. M. (2002). The Aggression Questionnaire: A validation study in student samples. *The Spanish Journal of Psychology*, 5, 45-53. doi:10.1017/s1138741600005825
- Geniole, S. N., Denson, T. F., Dixon, B. J., Carré, J. M., & McCormick, C. M. (2015). Evidence from Meta-Analyses of the Facial Width-to-Height Ratio as an Evolved Cue of Threat. *PLoS ONE*, 10(7), e0132726. doi:10.1371/journal.pone.0132726
- Giancola, P. R., & Chermack, S. T. (1998). Construct validity of laboratory aggression paradigms. *Aggression and Violent Behavior*, 3(3), 237–253. doi:10.1016/s1359-1789(97)00004-9
- Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Golomb, B. A., Cortez-Perez, M., Jaworski, B. A., Mednick, S., & Dimsdale, J. (2007). Point Subtraction Aggression Paradigm: Validity of a Brief Schedule of Use. *Violence*, 22(1), 95–103. doi:10.1891/vv-v22i1a006
- Goodman, R. A., Mercy, J. A., Loya, F., Rosenberg, M. L., Smith, J. C., Allen, N. H., ... Kolts, R. (1986). Alcohol use and interpersonal violence: alcohol detected in homicide victims. *American Journal of Public Health*, 76(2), 144–149. doi:10.2105/ajph.76.2.144

- Hadfield, P., Hobbs, D., Lister, S., & Winlow, S. (2002). Location, Location, Location. *Criminal Justice Matters*, 47(1), 34–35. doi:10.1080/09627250208553379
- Hehman, E., Flake, J. K., & Freeman, J. B. (2015). Static and Dynamic Facial Cues Differentially Affect the Consistency of Social Evaluations. *Personality and Social Psychology Bulletin*, 41, 1123-1134. doi:10.1177/0146167215591495
- Hehman, E., Leitner, J. B., & Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology*, 49(4), 747-754. doi:10.1016/j.jesp.2013.02.015
- John, N. W. (2008). Design and implementation of medical training simulators. *Virtual Reality*, 12(4), 269-279.
- Karelaia, N., & Hogarth, R. M. (2008). Determinants of Linear Judgment: A Meta-Analysis of Lens Model Studies. *Psychological Bulletin*. 134(3), 404, doi:10.2139/ssrn.1002859
- Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of personality*, 64, 311-338.
- Kramer, D., Maas, J., Wingen, M., & Kunst, A. E. (2013). Neighbourhood safety and leisure-time physical activity among Dutch adults: a multilevel perspective. *International Journal of Behavioral Nutrition and Physical Activity*, 10. doi:10.1186/1479-5868-10-11
- Miles, H. C., Pop, S. R., Watt, S. J., Lawrence, G. P., & John, N. W. (2012). A review of virtual environments for training in ball sports. *Computers & Graphics*, 36(6), 714-726.

- Monin, B., & Oppenheimer, D. M. (2005). Correlated averages vs. averaged correlations: Demonstrating the warm glow heuristic beyond aggregation. *Social Cognition*, 23(3), 257-278.
- Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology*, 35, 136–164.
- Neisser, U. (1980). On “Social Knowing.” *Personality and Social Psychology Bulletin*, 6(4), 601–605. doi:10.1177/014616728064012
- O'Connor, D. B., Archer, J., & Wu, F. W. C. (2001). Measuring aggression: Self-reports, partner reports, and responses to provoking scenarios. *Aggressive Behaviour*, 27, 79–101. doi:10.1002/ab.2
- Ritter, D., & Eslea, M. (2005). Hot Sauce, toy guns, and graffiti: A critical account of current laboratory aggression paradigms. *Aggressive Behavior*, 31(5), 407–419. doi:10.1002/ab.20066
- Sripada, C. S., Angstadt, M., McNamara, P., King, A. C., & Phan, K. L. (2011). Effects of alcohol on brain responses to social signals of threat in humans. *NeuroImage*, 55(1), 371–380. doi:10.1016/j.neuroimage.2010.11.062
- Taylor, A., & Barton, J. (2015). The detection of fearful and angry expressions in visual search. *Journal of Vision*, 15(12), 1354. doi:10.1167/15.12.1354
- Taylor, C. (1995). *Philosophical arguments*. Harvard University Press.
- Tedeschi, J. T., & Quigley, B. M. (1996). Limitations of laboratory paradigms for studying aggression. *Aggression and Violent Behavior*, 1(2), 163–177. doi:10.1016/1359-1789(95)00014-3

- Thoresen, J. C., Vuong, Q. C., & Atkinson, A. P. (2012). First impressions: Gait cues drive reliable trait judgements. *Cognition*, 124(3), 261–271.
doi:10.1016/j.cognition.2012.05.018
- Troscianko, T., Holmes, A., Stillman, J., Mirmehdi, M., Wright, D., & Wilson, A. (2004). What happens next? The predictability of natural behaviour viewed through CCTV cameras. *Perception*, 33(1), 87–101. doi:10.1068/p3402
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1130
- Ullman, S. E., & Brecklin, L. R. (2000). Alcohol and Adult Sexual Assault in a National Sample of Women. *Journal of Substance Abuse*, 11(4), 405–420. doi:10.1016/s0899-3289(00)00036-5
- Wilson, T. D., Houston, C., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, 4, 387–402

Appendix A. Ethical Approval for the Research in Chapters 2 & 3



Faculty of Science
University of Portsmouth
St Michael's Building
White Swan Road
PORTSMOUTH
PO1 2DT

Date 01/04/14

FAVOURABLE OPINION

Protocol Title: The Importance of Dynamic Body Stimuli in Threat Perception Research

Dear Liam,

Thank you for resubmitting your revised protocol for ethical review and for the clarifications provided.

Your responses have been reviewed and I am pleased to inform you that your application has been given a favourable opinion by the Science Faculty Ethics Committee. Thus, no further action is required on your part.

Please notify us in the future of any substantial amendments that may be required and send us a final study report.

Good luck with the study.

Best wishes,

A handwritten signature in black ink, appearing to read 'J. Kaminski'.

Dr. Juliane Kaminski
(Dept) Science Faculty Ethics Committee

CC -
Dr Chris Markham – Chair of SFEC
Sci.fac@port.ac.uk
Holly Shawyer – Faculty Administrator

Appendix B. Ethical Approval for the Research in Chapter 4



Mr Liam Satchell
Doctoral Student
Department of Psychology
University of Portsmouth

Liam.Satchell@port.ac.uk

Science Faculty Ethics Committee

Science Faculty Office
University of Portsmouth
St Michael's Building
White Swan Road
PORTSMOUTH
PO1 2DT

T: 023 9284 3379

ethics-sci@port.ac.uk

22 July 2015

FAVOURABLE ETHICAL OPINION

Study Title: **Body Processing in Threat Judgements**

Reference Number: **SFEC 2015-041** (Please quote this in any correspondence)

Thank you for resubmitting your application to the Science Faculty Ethics Committee (SEFC) for ethical review following the 1st SFEC review dated 7 July 2015 in accordance with current procedures¹.

I am pleased to inform you that SFEC was content to grant a favourable ethical opinion of the above research on the basis described in the submitted documents listed at Annex A, and subject to standard general conditions².

Please note that the favourable opinion of SFEC does not grant permission or approval to undertake the research. Management permission or approval must be obtained from any host organisation, including the University of Portsmouth or supervisor, prior to the start of the study.

Wishing you every success in your research

Yours sincerely,

A handwritten signature in black ink, appearing to read 'Jim House'.

Dr Jim House
Vice-Chair Science Faculty Ethics Committee

Information:

Claire Nee - PhD Supervisor

Paul Morris - PhD Supervisor

Lucy Akehurst - PhD Supervisor

Holly Shawyer - Faculty Administrator

¹ Procedures for Ethical Review, Science Faculty Ethics Committee, University of Portsmouth, October 2012 (to be updated).

² After ethical review – Guidance for researchers (Please read).

Appendix C. Ethical Approval for the Research in Chapter 5.



Faculty of Science
University of Portsmouth
St Michael's Building
White Swan Road
PORTSMOUTH
PO1 2DT

Liam.Satchell@port.ac.uk
24/11/14

Science Faculty Ethics Committee

Protocol Title: What's in a Walk? Perceptions of Motion and Morphology and Age, SFEC 2014-089
Date application received: 10/11/14
Date Reviewed: 24/11/14

FAVOURABLE OPINION WITH MINOR CONDITIONS

Dear Mr Satchell,

Thank you for your submission for ethical review. Having completed their review, members of the Science Faculty Ethics Committee have reached a Favourable opinion, with minor conditions, of your proposed research. Please note that you are *not* required to resubmit your documents confirming that these conditions have been actioned.

- Please amend the debrief sheet to show that the process for any potential complaint is Principal Investigator - Head of Department - UoP Complaints Officer.

Please notify the committee of any substantial amendments to the proposed procedures, send an annual report to the committee regarding study progress and a final study report once the study has concluded. Please send these to sci.fac@port.ac.uk.

Thank you for your submission and the Committee wishes you well with your study.

Dr Chris Markham – Chair of SFEC

A handwritten signature in black ink, appearing to read 'Chris Markham'.

CC -
Holly Shawyer – Faculty Administrator

If you would like to offer any feedback on the Science Faculty Ethics Committee process please email sci.fac@port.ac.uk to be forwarded to the Chair



Faculty of Science
University of Portsmouth
St Michael's Building
White Swan Road
PORTSMOUTH
PO1 2DT

Date 03/07/14

FAVOURABLE OPINION

Protocol Title: Whats in a walk? child's perception of motion and morphology

Dear Liam,

Thank you for resubmitting your revised protocol for ethical review and for the clarifications provided.

Your responses have been reviewed and I am pleased to inform you that your application has been given a favourable opinion by the Science Faculty Ethics Committee. Thus, no further action is required on your part.

Please notify us in the future of any substantial amendments that may be required and send us a final study report.

Good luck with the study.

Best wishes,

A handwritten signature in black ink, appearing to be 'J. Kaminski'.

Dr. Juliane Kaminski
(Dept) Science Faculty Ethics Committee

CC -
Dr Chris Markham – Chair of SFEC
Sci.fac@port.ac.uk
Holly Shawyer – Faculty Administrator

Appendix D. Ethical Approval for the Research Presented in Chapter 6.



Faculty of Science
University of Portsmouth
St Michael's Building
White Swan Road
PORTSMOUTH
PO1 2DT

Liam.Satchell@port.ac.uk
17/12/14

Science Faculty Ethics Committee

Protocol Title: SFEC 2014 – 040 C, What's in a Walk? Child Perceptions of Motion and Morphology
Date received PI response from Provisional Opinion Letter: 09/12/14
Date Reviewed: 17/12/14

FAVOURABLE OPINION SFEC 2014 – 040 C

Dear Mr Satchell,

Thank you for your submission for ethical review. Having completed their review, members of the Science Faculty Ethics Committee have reached a Favourable opinion of your proposed research.

Please notify the committee of any substantial amendments to the proposed procedures, send an annual report to the committee regarding study progress and a final study report once the study has concluded. Please send these to sci.fac@port.ac.uk.

Thank you for your submission and the Committee wishes you well with your study.

Dr Chris Markham – Chair of SFEC

A handwritten signature in black ink, appearing to read 'Chris Markham'.

CC -
Holly Sawyer – Faculty Administrator

If you would like to offer any feedback on the Science Faculty Ethics Committee process please email sci.fac@port.ac.uk to be forwarded to the Chair

Appendix E. Ethical Approval for the Research Presented in Chapter 7.



Faculty of Science
University of Portsmouth
St Michael's Building
White Swan Road
PORTSMOUTH
PO1 2DT

Liam.satchell@port.ac.uk
11/06/2015

Science Faculty Ethics Committee

Protocol Title: SFEC 2015 – 033, they're behind you! Threat judgements and memory of strangers
approaching from behind

Date received: 03/06/2015

Date Reviewed: 15/06/2015

FAVOURABLE OPINION - SFEC 2015-033

Dear Mr Satchell,

Thank you for your submission for ethical review. Having completed their review, members of the Science Faculty Ethics Committee have reached a Favourable opinion of your proposed research.

Please notify the committee of any substantial amendments to the proposed procedures, send an annual report to the committee regarding study progress and a final study report once the study has concluded. Please send these to ethics-sci@port.ac.uk.

Thank you for your submission and the Committee wish you well with your study.

A handwritten signature in black ink, appearing to read 'Chris Markham'.

Dr Chris Markham – Chair of SFEC

CC -

Holly Shawyer – Faculty Administrator

If you would like to offer any feedback on the Science Faculty Ethics Committee process please email ethics-sci@port.ac.uk to be forwarded to the Chair

Appendix F. The Buss Perry Aggression Questionnaire Short Form (Bryant and Smith, 2001).

Bryant and Smith refine the Buss-Perry Aggression Questionnaire (Buss & Perry, 1992) into a short form measure that is more reliable than the long form measure (by removing superfluous items). The physical aggression subscale has internal reliability of between $\alpha=.84$ and $\alpha=.86$ across three samples used by Bryant and Smith (2001). Using 5 samples, Webster et al. (2013) showed that the Buss-Perry Aggression Questionnaire – Short Form is internally reliable, shows strong test-retest reliability and predicts behavioural measures of aggression.

The scale (with standardised instructions) is presented on the next page. The items in **bold** are the Physical Aggression subscale items used in Chapters 3-7 (these were not bold when presented to participants).

References

- Bryant, F. B., & Smith, B. D. (2001). Refining the architecture of aggression: a measurement model for the Buss–Perry Aggression Questionnaire. *Journal of Research in Personality*, 35(2), 138-167.
- Webster, G. D., DeWall, C. N., Pond, R. S., Deckman, T., Jonason, P. K., Le, B. M., ... Bator, R. J. (2013). The brief aggression questionnaire: psychometric and behavioral evidence for an efficient measure of trait aggression. *Aggressive Behavior*, 40(2), 120–139. doi:10.1002/ab.21507

Instructions:

Please rate each of the following items in terms of how characteristic they are of you. Use the following scale for answering these items.

1	2	3	4	5	6	7
Extremely Uncharacteristic of me					Extremely Characteristic of me	

- **Given enough provocation, I may hit another person**
- **There are people who pushed me so far that we came to blows.**
- **I have threatened people I know.**
- I often find myself disagreeing with people.
- I can't help getting into arguments when people disagree with me.
- My friends say that I'm somewhat argumentative.
- I flare up quickly but get over it quickly.
- Sometimes I fly off the handle for no good reason
- I have trouble controlling my temper
- At times I feel I have gotten a raw deal out of life.
- Other people always seem to get the breaks.
- I wonder why sometimes I feel so bitter about things.

Appendix G. A discussion on reporting multiple correlation values.

Both examiners of this thesis raise an important question about my techniques of presenting results. How psychologists analyse and present data is a contemporary issue in psychology and a debate I welcome the opportunity to discuss in this discursive appendix.

Examiners' comment:

Please include a short section in the Appendices (a page or so) which explains your reasoning behind using Ferguson's (2009) criteria to make sense of multiple correlations rather than other potential approaches (i.e. significance testing with alpha adjustment for multiple tests; canonical correlation, etc.).

The use of Ferguson's criteria to make sense of correlations. In chapters two and four, I opted not to report p values in the tables of correlations (ρ and r respectively). Instead, I guide the reader to attend to the correlations of a 'decent' effect size. The problem with this method is attempting to define a 'decent' effect size. It is difficult for an author to put forward such a definition as there is no literature-standard assessment of an effect size, nor a standard arbitrary 'cut off' for meaningful effect (unlike the significance cut-off of .05 which has become a standard expectancy in psychology). To provide precedent for reviewers, readers and examiners of the paper, I cite Ferguson's (2009) published precedent of 'recommended minimum effect' (RME) sizes. Ferguson also expands the RMEs to define 'moderate' and 'large' effects in his terms (which I use in Chapter 4 for a more conservative analysis of the correlation values). Ferguson is used here to provide an established and consistent evaluation of the size of an effect (considering the issues with significance testing, see below). My personal evaluation of an effect size, which I, like most researchers, primarily use (until there is wider consistency in the literature) is that correlations should be considered notable (albeit small) when they are larger than .30. I would consider a correlation of .40 or larger to be important, especially in a social judgment context, where the literature typically contains small effect sizes (Kenny, Albright, Malloy & Kashy, 1994; Richard, Bond & Stokes-Zoota, 2003). I generally agree with Ferguson's (2009) interpretation of effect sizes. Rather than describing the data with my own points of view, which is typical of the literature, I cite a peer reviewed journal article as justification for my interpretation.

More generally, I am against cut-off points in statistical testing. As Fidler, Geoff, Mark and Neil (2004) note "the dichotomous accept/rejection decision outcome of NHST provides only the illusion of objectivity", (p. 619). The issues with deciding on an arbitrary p value equally apply to effect size cut-offs. It would be a positive move for researchers to simply report on the nature of their data by reporting effect sizes and allowing readers to interpret the size themselves. However, for ease of reading and to address the varied statistical knowledge of potential readers, I opt to include cut-offs for clarity in this thesis.

Why not use significance testing with a corrected alpha? There is an increasing pressure in the psychological literature to move away from using significance testing altogether. For example, I have recently been instructed to remove p values by an editor in the revision process of a journal article (Satchell & Pearson, under revisions). The arguments against significance testing are hardly new. Researchers in the 1960s (e.g. Meehl, 1967) were extremely critical of null hypothesis significance testing (NHST). The concern with interpreting and publishing with multiple corrected p values has also been specifically addressed in the literature: “Ironically [problems with publication issues around p values] can be exacerbated by the use of statistical procedures designed to correct for multiple applications of a test statistic” (Nickerson, 2000, p272). There are many critiques of p values, but the critique I most closely follow and is most relevant to the question of multiple corrections is a philosophical one (highlighted in Cohen, 1994).

It is important to note that correlation coefficients describe relationships (the computed correlation values are entirely based on the data present in the samples) and p values are inferential (they add interpretation). This inference is, “Given that H_0 is true [in the whole population], what is the probability of these data [occurring]?” (Cohen, 1994, p.997). This argument is designed to relate the data here to the wider population. This inferential p value is derived from the descriptive correlation value (be it r or ρ) and the sample size. It is used to create a judgment of the significance of the data when compared to the arbitrary value of $p=.05$. It is often considered good practice to ‘correct’ a p value cut-off when presenting multiple tests due to the heightened probability of Type I error (finding an effect where there is none) by chance. This is done by increasing the assumed probability that the statistic derived from the data is due to chance. However, given that our primary logic argument behind NHST is bound by a question of the significance for this *current* test statistic to the wider population (H_0 is true in the wider population), then correcting a p value is changing our original premise. I would suggest that multiple corrections lead to a logic argument that is fundamentally unwieldy and will demonstrate it with this following example:

If the relationship between variables A and Z are demonstrated using a computed correlation value of $r=.38$, given the sample I have here, I could assign a value of significance to this relationship. That is, if I have tested 45 people, I could assume that these is a 1% chance of this effect being spurious when applied to the wider population, thus $p=.01$ (from the standard critical values of r table). This result is now deemed significant and noteworthy.

Given these data, we can assume there is a 1% chance that H_0 is true in the overall population. However, a more rigorous study would also point out that feature A of a person only represents a subset of that person’s behaviour and investigating the effect of behaviours B to M when correlated with Z could also show meaningful patterns. In fact, Z is not the only interesting possible outcome and N to Y also could correlate interestingly with variables A to M to show a more holistic network of relationships (such as showing how various aspects of gait relate to various aspects of personality.) Now when correlating A to M with N to Z , we no longer consider $r_{A|Z}=.38$ meaningful because we have 169 total

correlations and have corrected our alpha to a more conservative value (say $p < .001$ or $p < .0003$). By trying to more holistically describe the wider relationships in the data set, (for example $r_{B|Y} = .38$ as well) we now consider none of these findings noteworthy, despite the fact these are two independent tests (i.e. if A to M are orthogonally derived measures, there is no reason to assume a relationship between A and B .)

The practice of multiple corrections creates an unwieldy logical argument because it is based on the premise that running more than one test (or two independent tests) increases the likelihood of accepting a null hypothesis. The logical argument then becomes: *given that H_0 is true and no other tests have been run what is the probability of these data occurring?* Given that not all possible tests on a data set are ever run and presented, the simplest way to address multiple correction errors is to report fewer tests.

It is certainly true that an increase in number of tests conducted increases the likelihood of Type I error, as well as a false assumption that there is an effect in the whole population from the subset of the population that was tested. However, this ‘error’ in finding an effect is derived from those descriptive r or rho values (which are entirely based on the data collected). Type I error, in this sense, is but a demonstration of an effect within the specific sample. The issue of spurious findings, or error, is eliminated by binding the results of the research to the currently studied population; “There was a moderate correlation between A and Z in our sample”.

There is a general awareness in research about methodological representativeness (i.e. the sample is 80% female so how can we generalise the findings?) and these arguments are frequently raised in discussion sections and reviews of others’ works. However, the issue of generalisability of findings can only be addressed with more general samples and thorough replication (including sampling of targets, see Monin & Oppenheimer, 2014; Westfall, Judd & Kenny, 2015). Psychologists cannot, and should not, infer whole population behaviours from subsets and researchers do report this; either implicitly or explicitly. In short, researchers do not conduct psychological research and then claim the findings explain all humankind’s behaviour. Research is moving away from p values in general due to issues with interpretation (and the influence p values have on publication biases.) While I do report p values in many of the studies in this thesis, this is done to improve the likelihood of publication (in fact the published version of Chapter 2 contains [uncorrected] p values after the editor’s instruction); however, I draw a clear line with multiple comparisons. It distorts the story the data tells by relying on a p value that is more influenced by the number of tests I choose to report.

Why not use canonical correlations? Canonical correlations are used to reduce the number of tests being run on a dataset. It is similar to factor analysis (reducing dependent measures, K) or a cluster analysis (reducing sample size, N) and reveals the fewest number of correlations that show unique variance (typically canonical correlations demonstrate an orthogonal solution). Canonical correlations are a technique used to decrease the number of tests so to avoid multiple correction issues. Thus, canonical correlations are not necessary if the above critique of multiple corrections is valid. It is

more descriptive and representative of the data to include a full explanation of the data analysis, rather than reducing the number of factors being analysed. Canonical correlations show how two categories of variables share variance; however, they are best applied when trying to explain variance of subsets of measures rather than two groups of separate measures. In this Thesis, demonstrating that the ‘sub measures’ of the two dimensions of variables could be correlated (personality traits and gait in Chapter 2, and time and accuracy-observation preference values in Chapter 4) is no more informative when reduced to fewer factors. In fact, more information is provided when they are reported as separate descriptive. Furthermore, canonical correlations are a parametric analysis, which is not suitable for the data I treat non-parametrically in Chapter 2.

In summary. I thank the examiners for allowing me to discuss my choice of analytical style. This thesis reports significance testing as it is still typical to present *p* values and in most cases there is no consequence of reporting in more information. However, multiple corrections would misrepresent the overall analyses and I opt to report uncorrected test statistics in my results sections.

References

- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49(12), 997–1003. doi:10.1037/0003-066x.49.12.997
- Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), 532.
- Fidler, F., Geoff, C., Mark, B., & Neil, T. (2004). Statistical reform in medicine, psychology and ecology. *The Journal of Socio-Economics*, 33(5), 615–630. doi:10.1016/j.socec.2004.09.035
- Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: Acquaintance and the big five. *Psychological Bulletin*, 116(2), 245–258. doi:10.1037/0033-2909.116.2.245
- Meehl, P. E. (1967). Theory-Testing in Psychology and Physics: A Methodological Paradox. *Philosophy of Science*, 34(2), 103–115. doi:10.1086/288135
- Monin, B., & Oppenheimer, D. M. (2014). The limits of direct replications and the virtues of stimulus sampling. *Social Psychology*, 45, 299–300.
- Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*, 5(2), 241–301. doi:10.1037/1082-989x.5.2.241
- Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One Hundred Years of Social Psychology Quantitatively Described. *Review of General Psychology*, 7(4), 331–363. doi:10.1037/1089-2680.7.4.331

Satchell, L., Morris, P., Mills, C., O'Reilly, L., Marshman, P., & Akehurst, L. (2016). Evidence of Big Five and Aggressive Personalities in Gait Biomechanics. *Journal of Nonverbal Behavior*. doi:10.1007/s10919-016-0240-1

Satchell, L. & Pearson, D. (under revisions) The Social Salience of Student's Sub-clinical Psychopathic Personality. *The Journal of General Psychology*.

Westfall, J., Judd, C. M., & Kenny, D. A. (2015). Replicating Studies in Which Samples of Participants Respond to Samples of Stimuli. *Perspectives on Psychological Science* 10(3) 390-399. doi:10.2139/ssrn.2449567

FORM UPR16

Research Ethics Review Checklist

Please include this completed form as an appendix to your thesis (see the Postgraduate Research Student Handbook for more information)

Postgraduate Research Student (PGRS) Information		Student ID:	491110
PGRS Name:	Liam Satchell		
Department:	Psychology	First Supervisor:	Paul Morris
Start Date: (or progression date for Prof Doc students)	September 2013		
Study Mode and Route:	Part-time <input type="checkbox"/> Full-time <input checked="" type="checkbox"/>	MPhil <input type="checkbox"/> PhD <input checked="" type="checkbox"/>	MD <input type="checkbox"/> Professional Doctorate <input checked="" type="checkbox"/>

Title of Thesis:	Factors Affecting the Accurate Detection of Trait Aggression from Gait
Thesis Word Count: (excluding ancillary data)	44,353

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University's Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study

Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

UKRIO Finished Research Checklist:

(If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee rep or see the online version of the full checklist at: <http://www.ukrio.org/what-we-do/code-of-practice-for-research/>)

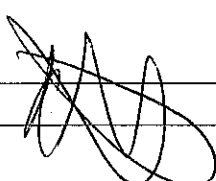
a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>
b) Have all contributions to knowledge been acknowledged?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>
c) Have you complied with all agreements relating to intellectual property, publication and authorship?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>
d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>
e) Does your research comply with all legal, ethical, and contractual requirements?	YES <input checked="" type="checkbox"/> NO <input type="checkbox"/>

Candidate Statement:

I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s)

Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCREC):	SFEC(2104-089), SFEC(2015-041), SFEC(2015-033)
--	--

If you have *not* submitted your work for ethical review, and/or you have answered 'No' to one or more of questions a) to e), please explain below why this is so:

Signed (PGRS):		Date: 31/3/2016
----------------	---	-----------------